

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 13:34:09 ; Search time 7963.34 Seconds  
(without alignments)  
4867.828 Million cell updates/sec

Title: US-10-051-909-32  
Perfect score: 4075  
Sequence: 1 IRSGSWLAVQTPTPDLRR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074037\_14048/app\_query.fasta\_1.1678  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909@CGN\_1\_1\_9091@runat\_13042005\_074037\_14048 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3517	86.3	2824	6 AR208565	AR208565 Sequence
2	3395	83.3	2934	8 AK120560	AK120560 Oryza sat
3	3155	77.4	2665	8 AY165599	AY165599 Saccharum
4	3097	76.0	2378	8 AK099716	AK099716 Oryza sat

ALIGNMENTS

RESULT 1	AR208565	LOCUS	AR208565	Sequence 1 from patent US 6383776.	2824 bp	DNA	linear	PAT 20-JUN-2002
DEFINITION	AR208565	ACCESSION	AR208565	VERSION	AR208565.1	GI:21509752		
KEYWORDS		SOURCE	Unknown.	ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 2824)	AUTHORS	Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.	TITLE	Plant sugar transport proteins			
JOURNAL	Patent: US 6383776-A 1 07-MAY-2002;	FEATUES	Location/Qualifiers	source	1..2824	/organism="unknown"	/mol_type="unassigned DNA"	

Alignment Scores:								
Pred. No.:	3.94e-225	Length:	2824					
Score:	3517.00	Matches:	684					
Percent Similarity:	95.18%	Conservative:	27					
Best Local Similarity:	91.57%	Mismatches:	34					
Query Match:	86.31%	Indels:	2					
DB:	6	Gaps:	2					

US-10-051-909-32 (1-800) x AR208565 (1-2824)			
Qy	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr	75
Db	238	ATGGGGGCGCGTGTATGTCGCCATCGCGGCTTATCGGCAACTTGCTGCAGGGCTGG	297
Qy	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn	95
Db	298	GACAATGCGACAATTGCTGGAGCCGTCTGTATACATAAAGAAGGAATTCACCTGCAGAGC	357
Qy	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	358	GAGCCTCTGATCGAAGGCCTCATCGTCGCCATGTTCTCATTTGGGGCAACAGTCATCACA	417
Qy	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135
Db	418	ACATCTCCGGGGCCAAAGGCTGACTGCGTTGGTAGAGGCCCATGCTGGTCCGCTCGGCT	477
Qy	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu	155
Db	478	GTCTCTACTTCGTACAGTGGCTGGTGATGCTTTGGGCGCCAATTGTGTACATCTTGCTC	537
Qy	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175
Db	538	CTCGCAAGGCTCATTTGATGGGTTCCGTATCGGTTTGGCGGTACACATTGTTCTCTCTAC	597
Qy	176	IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe	194
Db	598	ATCTCCGAAACTGCACCGCACAGANATTCTTGGGGCTGNTNGAACACGTTGCCGAGTTC	657
Qy	195	serGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer	213
Db	658	ATTGGGTCAGNGGAGGATGTTCTCTCTACTGATGATGTTTGGGATGTCCTCATG	717
Qy	214	ProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePhe	233
Db	718	CCCAAACTGATTGGAGGCTCATGCTTGGAGTTCTGTCTGATCCCGTCACTATNTACTTT	777
Qy	234	GlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMet	253
Db	778	GGACTGACTGCTTCTACTTGGCTGAATCACCAAGTGGCTTGTNAGCAAGGAAGGATG	837
Qy	254	AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu	273
Db	838	GCGGAGCGAAGAGAGTGTGTCAAAAGGCTGCGGGGAAGAGAAGATGTCTCANGGGGANG	897
Qy	274	SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle	293
Db	898	GCTCTTCTAGTTGAAGGTTTGGGGTCCGTAAGATAACGTAATTTNAGAGTACATCAT	957
Qy	294	GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr	313
Db	958	GGACCTGCCACCGGCGAGCCGATGATCTTGTAACGTGACCGTGATAAGGAACAAATCACA	1017
Qy	314	LeuTyrGlyProGluGluGlyGlnSerTyrIleAlaArgProSerLysGlyProIleMet	333
Db	1018	CTTTATGGGCTGAAGAAGGCCAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATG	1077
Qy	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
Db	1078	CTTGGAAAGTGCTTTCTCTTGCACTCTCGTCATGGGAGCATGGTGAACCAGAGTGTACCC	1137
Qy	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
Db	1138	CTTATGGATCCGATTGTGACACTTTTGGTAGTGTCATGAGAATATGCCTCAAGCTGGA	1197
Qy	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1198	GGAAGTATGAGGAGCACATTGTTTCCAAACTTTGGAAAGTATGTTTCAGTGTACAGATCAG	1257
Qy	394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
Db	1258	CATGCCAAAAAATGAGCAGTGGGATGAAGAGAAATCTTTCATAGGGATGACGAGGATACGCA	1317
Qy	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db	1318	TCTGATGTCAGGAGGTGACTATGAGGACAATCTCCATAGCCCATTTGCTGTCCAGGCAG	1377
Qy	434	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer	453
Db	1378	GCAACAGGTGCGGAAGGAAGGACATTGTGCACCATGTCACCGTGGAAGTCTTTGAGC	1437
Qy	454	MetArgArgGlnThrLeuLeuGlyGlyGlyGlyAspGlyValSerSerThrAspIleGly	473
Db	1438	ATGAGAAGGCAAGCCTCTTAGGGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGGT	1497
Qy	474	GlyGlyTyrGlnLeuAlaTrpLysTyrSerGlyLysGluGlyAsnGlyArgLysGlu	493
Db	1498	GGGGGATGGCAGCTTCTTGGAAATGGTCAGAGAAGGAAGGTGAGAATGGTAGAAAGGAA	1557
Qy	494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer	513
Db	1558	GGTGGTTCAAAAAGAGTCTACTTGCACCAAGAGGAGTTCCTGGCTCAAGAAGGGGCTCA	1617
Qy	514	IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533
Db	1618	ATTGTTTCACCTTCCCGGTGGTGGCGATGTTCTTGGGGTAGTGAGTTTGACATGCTGCT	1677
Qy	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553
Db	1678	GCCTTAGTAAGTCAGTCAGCACTTTTCTCAAAGGGTCTTGTGAAACCAACGATGTTCAGAT	1737
Qy	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPhe	573
Db	1738	GCTGCCATGGTTCACCCATCTGAGGTAGCTGCCAAAGGTTTCACGTTGGAAAAGATTGTTT	1797
Qy	574	GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe	593
Db	1798	GAACCTGGAGTGAGCGCTGCCCTGTTAGTCGGTGTGGAATTTCAGATCCTTCAACAGATT	1857
Qy	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla	613
Db	1858	GCTGGAATAAACCGTGTTCTGTAATAACCCCAACAAATCTTGAGCAAGCTGGTGTGGCA	1917
Qy	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu	633
Db	1918	GTTATTCTTTCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGTATCAGTTCTCTC	1977
Qy	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653
Db	1978	ACTACCTTACTAATGCTTCCTTGCAATGGCTTTGCCATGCTGCTTAAGGATCTTCCGGA	2037
Qy	654	ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal	673
Db	2038	AGAAGGTTTTTGTGCTGCTAGGCACAATTCCAATCTTGATAGCATCTCTAGTTATCTGGTT	2097
Qy	674	ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal	693
Db	2098	GTGTCCAATCTAATTGATTGGGTACACTAGCCCCATGCTTTGCTCTCCACCATCAGTGTT	2157
Qy	694	IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla	713
Db	2158	ATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA	2217
Qy	714	GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr	733
Db	2218	GAGATCTTCCAAACCAGGGTTCGTGGCCTCTGTATTGCCATTGTGCCCTTTACATTTCTGG	2277
Qy	734	IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla	753
Db	2278	ATCGGAGATATCATCGTCACCTACAGCCCTTCTCTGTGATGCTGAATGCTATTGGAGCTGGCG	2337
Qy	754	GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys	773
Db	2338	GGTGTTTTTCAGCATATATGCAGTCGTATGCTTGATTTCCTTTGTGTTTCGTTCTCTTAAG	2397





Db	334	CTGCTGCAGGGATGGGACAAATGCCACCATTCAGGTGCTGTACTATACATAAAGAGGAA	393
QY	91	PheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGly	110
Db	394	TTCAAGCTAGAAAGTGAAGCCCACTGTGGAGGGGCTAATCGTGGCCATGTCTACTGATGGT	453
QY	111	AlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMet	130
Db	454	GCAACCATCATCAACAACATCTCAGGGCCGGTATCAGACTGGATCGGCCGGCGCCTATG	513
QY	131	LeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsn	150
Db	514	CTCATTTCTCTCAATTTCTTACTTCTCAGCAGCCTCATCATGTTATGGTCCCCGAAT	573
QY	151	ValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr	170
Db	574	GTCTATGTGTACTGCTGCACGCCTTATAGATGGATTGGCATCGGCTTGGCTGTCTACA	633
QY	171	LeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr	190
Db	634	CTTGTACCTTTGTACATCTCAGAGACAGCTCCTTCAGAGATCAGGGGTTTGTGAATACA	693
QY	191	LeuProGlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMet	210
Db	694	CTGCCACAGTTCAAGGGTTCAGGAGGGATGTTCTTGTCTACTGTCATGGTGTGGGATG	753
QY	211	SerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeu	230
Db	754	TCACTGTTCGCGTCACTGACTGGAGAAATTATGCTTGGTGTCTCGCAATCCCTTCGTTG	813
QY	231	PhePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLys	250
Db	814	TTTTTCTTCGGATTGACAAATATTCTATCTGCCAGAATCACCAAGATGGCTTGTTCAGCAA	873
QY	251	GlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSer	270
Db	874	GGGCGGATGGCTGAGGCAAGAAGGTATTACAAAATTACGTGGGAGAGAGGATGTCTCA	933
QY	271	GlyGluLeuSerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlu	290
Db	934	GGAGAAATGGCTCTTGTGTGAAGTTTGGAGTTTGGAGCTGACACTTCCATTGAAGAG	993
QY	291	TyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGlu	310
Db	994	TACATCATCGGACCGGCTATAGAGCCAGCTGTATGATGCATGTTGTTGATGGCGATAAGAC	1053
QY	311	GlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGly	330
Db	1054	CAAATAACACTGTACGGGCTGTAGAGGGCCAATCATGGATTGCTCGACCTTCCAGGGA	1113
QY	331	ProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGln	350
Db	1114	CCCAGCATTTCTTGAAGTGTGCTTCTCTTACATCTCGTCATGGCAGCATGGTGAACACAG	1173
QY	351	SerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetPro	370
Db	1174	AGCGTGGCACCTTATGGATCCTATAGTCAACGCTTTTCGGCAGTGTCCATGAGAACATGCCT	1233
QY	371	GlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerVal	390
Db	1234	CATGCTGGAGGAAGCATGCGAGTACATTGTTTCCCAACTTGGTAGTATGTTTAGTGTG	1293
QY	391	ThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGlu	410
Db	1294	ACAGATCAGCACCCCCAAGTTGATCAATGGGATGAGGAGAACCTTCATAGGGATGATGAG	1353
QY	411	GluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeu	430
Db	1354	GAATACGCGTCTGATGGTGTGGAGGTGACTATGAAGACAAATGTCCACAGCCCCGTTGCTG	1413
QY	431	SerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySer	450

D	b	1414	TCCCGACAGACCACGAGCGCAGAAGGGGAAGGACATTTGCACACCATGCTCATCTGTGGAAGT	1473
Q	y	451	AlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThr	470
D	b	1474	GCTCTGAGTATGAGAAAGAACGCCCTCTTGAAAGAGGGTGGCGGCGGTGAGCAGCACT	1533
Q	y	471	AspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGly	490
D	b	1534	GGCATTGGTGGGGATGGCAGCTTGTCATGGAAATGGTCAGAGCGAAGGCGAGGATGGT	1593
Q	y	491	ArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArg	510
D	b	1594	AAGAAAGGAAGGAGGATTTAAAAGGATCTACTTGCACCAAGAGGAAGTGCAGGATCAAGA	1653
Q	y	511	ArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheVal	530
D	b	1654	AGGGGCTCAGTTATTTTCACITTCCTGGTGGAGGGGACGCTCCTGAAGGCAGCGAATTCATA	1713
Q	y	531	HisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArg	550
D	b	1714	CATGCTGCCGCTTTGGTAAGCCAACGAGCACCTTTACTCCAAGGATATCATTTGAACAGCGT	1773
Q	y	551	MetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLys	570
D	b	1774	ATGTCCCGTCCAGCCATGATTTCATCCATCAGAGGCAGCTGCCAAAGGTTCAAGCTGGAAA	1833
Q	y	571	AspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeu	590
D	b	1834	GATTTGTTTGAAACCTGGAGTAGGCGTGCGTTGTTAGTCGGTGTGGGAATTCAAATCCTT	1893
Q	y	591	GlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAla	610
D	b	1894	CAACAGTTTGCTGGAAATAAATGGGGTTCTCTATTACACTCCACAATACTTGAGCAAGCG	1953
Q	y	611	GlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIle	630
D	b	1954	GGGGTGGCAGTTCTCCTTTCCAAATCTTGGCCTCAGTTTCAGCATCAGCTTCCATCTTGATC	2013
Q	y	631	SerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAsp	650
D	b	2014	AGTTCTTGACCACCTACTGATGCTTCTTAGCATTTGGTTTAGCCATGAGACTTATGGAC	2073
Q	y	651	LeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal	670
D	b	2074	ATCTCTGGAAGAAGGTTTTCTGTCTTCTGGGCACAAATTCAGTCTTGATAGCATCTCTAGTT	2133
Q	y	671	IleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThr	690
D	b	2134	GTCTTGGTTGTGTCCAATGTTATCGACCTGGGTACAGTGGCCCCACGCCGCACTCTCCACA	2193
Q	y	691	ValSerValIleValTyrPheCysCysPheValMetMetGlyPheGlyProIleProAsnIle	710
D	b	2194	ATCAGCGTCATCATCTACTTCTGTCTGCTTCGTTCATGGGATTCGGTCCGATCCCACCAACTT	2253
Q	y	711	LeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPhe	730
D	b	2254	CTGTGTGGGAGATCTTCCCACACTAGGGTCCCGGGCATCTGCATTGCCATCTGCGCCCTG	2313
Q	y	731	ThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIle	750
D	b	2314	ACATTCTGGATTGGTGATATCATAGTCACCTACAGCCTCCCTGTGATGCTGAATGCCATC	2373
Q	y	751	GlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheVal	770
D	b	2374	GGCCTAGCAGGTGCTTTTGGTATATACGCGAGTTGTTTGCTCGATTGCTTCGTTCGTTCGTC	2433
Q	y	771	PheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAla	790
D	b	2434	TTCTCTCAAGGTCCTCCGAGACGAAGGGCATGCCGCTCGAAGTCATCACCGAGTTCTTCGCT	2493
Q	y	791	ValGlyAlaLysGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla	798
D	b	2494	GTGGTGCAAGCAAATGCAGGCT	2517

RESULT 3  
AY165599  
LOCUS  
DEFINITION  
Saccharum hybrid cultivar putative sugar transporter type 2a mRNA, complete cds.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Saccharum hybrid cultivar  
Saccharum hybrid cultivar  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.  
REFERENCE  
1 (bases 1 to 2665)  
AUTHORS  
Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.  
TITLE  
Identification of a novel sugar transporter homologue strongly expressed in maturing stem vascular tissues of sugarcane by expressed sequence tag and microarray analysis  
JOURNAL  
Plant Mol. Biol. 52, 371-386 (2003)  
REFERENCE  
2 (bases 1 to 2665)  
AUTHORS  
Casu,R.E., Grof,C.P.L., Rae,A.L., McIntyre,C.L., Dimmock,C.M. and Manners,J.M.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (17-OCT-2002) Division of Plant Industry, Commonwealth Scientific and Industrial Organisation, 120 Meiers Rd, Indooroopilly, QLD 4068, Australia  
FEATURES  
source  
1..2665  
/organism="Saccharum hybrid cultivar"  
/mol\_type="mRNA"  
/db\_xref="taxon:128810"  
97..2334  
/codon\_start=1  
/product="putative sugar transporter type 2a"  
/protein\_id="AAO37640.1"  
/db\_xref="GI:31505504"  
/translation="MSGAALVAIVASIGNLLOQWDNATISAALLYIKKEFKLESEPTV EGLIVAMSLIGATIIITFGSPVSDWIGRLPMLILSSVLVFSLLIMLWSPNVYVLLLA RLINGFGVLAVTLPLYISETAPPEIRGLLNTLPQFSGSGGMFLSYCMVFGMSLLPS PDWRIMGLVALPSLFFFFGLTIFYLPESPRLVSKRWBAKKVLQRLRGKDDVTGEM ALLVEGLEVGDDTFIEEYIIGPAIEPADHDVADGRERIILFGPEEQSQSWIARPSKGS SMLGSVLSIGSRHGSVMVNQLPTDPIVTLFGSVHENTSLAGGSMRRIVPSKFDMSIS ITDQHPKTEQWDEENLHRDNEYATDGAASDYENNVSPLLRSQNTSAEKGKIAHHGH RGSALGLRRRSLSDGEGEAAITSTGIGGWQLAWKWSERGEDGKKEGSFKRIYILHOEG VAGSRRGSVSLPGGGDASEGKFIHAAALVQSALYPRDITEQRMAGPATMHPSEAA AKVPSWRDLFEPGVRALLVGIGIQLQFAGINGLVYTPQIMEQAGVAVLISNLGL SSASASILSSVTALLMLPSIGLAMRLMDVSGRRFLLSLTPVLIASLIVLVSVNVE LGTVVHAVLSTISVITYLCCFKMGFGPIPNILCAEFFFFRVRGICIAICALIFWVGDI IVTYSLPVMLNAIGLEGVFGIYAVACAIAFVFVYLKVPETKGMPLLEVITEFFAVGAKQ AVAKA"  
97..156  
/note="putative; transmembrane-region site"  
115..774  
/note="pfam00083; Region: sugar transporter"  
238..297  
/note="putative; transmembrane-region site"  
313..372  
/note="putative; transmembrane-region site"  
406..462  
/note="putative; transmembrane-region site"  
511..570  
/note="putative; transmembrane-region site"  
589..654  
/note="putative; transmembrane-region site"  
658..1794  
/note="Region: large cytoplasmic loop"  
1624..2298  
/note="pfam00083; Region: sugar transporter"  
1672..1731  
/note="putative; transmembrane-region site"

misc\_feature  
1813..1869  
/note="putative; transmembrane-region site"  
misc\_feature  
1897..1953  
/note="putative; transmembrane-region site"  
misc\_feature  
1969..2022  
/note="putative; transmembrane-region site"  
misc\_feature  
2089..2151  
/note="putative; transmembrane-region site"  
misc\_feature  
2191..2247  
/note="putative; transmembrane-region site"  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.35e-201 Length: 2665  
Score: 3155.00 Matches: 601  
Percent Similarity: 89.56% Conservative: 68  
Best Local Similarity: 80.46% Mismatches: 78  
Query Match: 77.42% Indels: 0  
DB: 8 Gaps: 0  
US-10-051-909-32 (1-800) x AY165599 (1-2665)  
QY 54 AspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGln 73  
Db 91 GAAACATGTCGGAGCTGCTCTGTGTCATAGTAGCTTCGATCGCAATCTGCTTCAG 150  
QY 74 GlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeu 93  
Db 151 GGGTGGACAAATGCCACCATCTCAGCTGCTCTATTATATATAAAGGAATTCAAAC 210  
QY 94 GlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIle 113  
Db 211 GAAAGTGAGCCTACTGTGGAGGGGCTGATTGTGGCCATGTCACTAATTGGAGCAACTATC 270  
QY 114 ValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeu 133  
Db 271 ATTACAACATTCTCTGGACCAGTATCAGACTGGATTGGCCGGCTCCCTATGCTCATCCTC 330  
QY 134 SerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrVal 153  
Db 331 TCCTTCAGTTCTGTACTTTGTTAGCAGCCTCATATTGCTATGTTCCCGAATGCTATGTA 390  
QY 154 LeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValPro 173  
Db 391 CTACTGCTGGCACGCTCTTATAAATGGATTCCGCTCGGCTTGGCTGCTACTCTTGTTCCT 450  
QY 174 LeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGln 193  
Db 451 CTGTACATTCAGAAACAGCTCCTCCAGAGATAAGAGGTTTGTGTAATACACTGCCACAG 510  
QY 194 PheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213  
Db 511 TTCAGTGGTCCAGGGGGGATGTTTTTGTGCTACTGTCATGGTGTTCGGGATGCTACTGCTG 570  
QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233  
Db 571 CCATCACTGATTGGAGGATTATGCTTGGGGTCTCGCGTTACCTTCATTGTTTCTTT 630  
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253  
Db 631 GGTTTGACAAATATTTTATCTGCCCTGAATCACCAGATGGCTTGTACGCAAGGGCGAATG 690  
QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273  
Db 691 GCAGAGGCAAAAGAGGTGTTGCAAGATTACGTGGGAGGATGACGTCACAGGCGAAATG 750  
QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293  
Db 751 GCTCTTCTTGTGAAGGGTTGGAGGTCGGAGGGGATACCTTCATTGAAGAGGTACATCATT 810  
QY 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313  
Db 811 GGTCCTGCTATCGAACCCAGCTGTATGATGACCACGTCGCTGATGGTGATAGAGAACGAATA 870



Qy	314	LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaAargProSerLysGlyProIleMet	333
Db	871	CTATTGGGCCTGAAGAGGGCCAATCATGGATAGCTGGCCTTCCAAGGGTCCAGCATG	930
Qy	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
Db	931	CTTGGAAGTGTCTTTCTATCGGATCTCGTCATGGCAGCATGGTGAACCAAAATTTACCC	990
Qy	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
Db	991	CITACGGATCCTATTGTACGCCTTTTGGGAGTGTCCATGAGAACATCTCTCTCGCTGGA	1050
Qy	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1051	GGAAGTATGAGGAGAATAGTGCCTTCTAAATTTGCAGTATGATCATGATCACAGATCAG	1110
Qy	394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
Db	1111	CACCCCAAGACCGAGCAGTGGGATGAGGAGAACCTCCATAGGGATAATGAGGAGTATGCA	1170
Qy	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db	1171	ACTGATGGTGTCTGCAAGTGACTATGAAACAATGTACACAGCCCTCTGCTGTCCGACAG	1230
Qy	434	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer	453
Db	1231	AATACAAGTCCCGAAGGGAAGGCATTGCACATCACGGGCACCGTGGAAAGTCTTTGGGT	1290
Qy	454	MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly	473
Db	1291	TTGAGAAGAAGAACCTGTCCGATGAGGGTGGTGAGGCAGCAACCAAGCTGGCATTGGT	1350
Qy	474	GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu	493
Db	1351	GGCGGATGGCAGCTCGCCTGGAAATGGTCAGAGCGAGAAGGTGAGGACGGTAAAGAAGAA	1410
Qy	494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer	513
Db	1411	GGGAGTTTCAAAAGAAATCTACTTACACCAAGAGGGAGTTGCTGGCTCAAGAAGGGGATCT	1470
Qy	514	IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533
Db	1471	GTTGCTCACTTCTCTGGTGGAGGTGATGCCTCTGAAGGTGGTAAAGTTTCATACATGCTGCT	1530
Qy	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553
Db	1531	GCTCTGGTCAGCCAGTCAGCTCTTTACCCGAGGGATATTACAGAACAGCGCATGGCCGGT	1590
Qy	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe	573
Db	1591	CCAGCTACAATGCACCCATCAGAGGCAGCTGCCAAAAGTTCCAAGCTGGAGAGACCTGTTT	1650
Qy	574	GluProGlyValArgArgAlaLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe	593
Db	1651	GAACCTGGTGTGAGACGTGCCCTGTTGGTTGGTATTGGAAATTCAGATCCTTCAACAGTTC	1710
Qy	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla	613
Db	1711	GCGGGAATAAATGGGGTTCTCTACTACACTCCGCAAATATTGGAGCAAGCTGGTGTGGCA	1770
Qy	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeu	633
Db	1771	GTTCTTATTCCAATCTTGGTCTCAGTTCCGGCATCAGCATCCATCCTTAATTAGTTCTGTC	1830
Qy	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653
Db	1831	ACTGCCCTACTTATGCTCCCTAGCATTTGGTTTAGCCCATGAGACTTATGGACGTATCTGGA	1890
Qy	654	ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal	673
Db	1891	AGAAGGTTTCTGCTGCTAAGCACAAATCCAGTCTTGATAGCATCTTTAATTGCTCCTGGTT	1950

Qy	674	ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal	693
Db	1951	GTGTCTAATGTTATCGAGTTGGGTACAGTGGTGCATGCTGTCTCTCCACAATTAGTGTC	2010
Qy	694	IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla	713
Db	2011	ATCACCTACTTATGCTGCTTCAAAATGGGCTTGGCCCCCATTTCCCAACATTTCTATGTGCA	2070
Qy	714	GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp	733
Db	2071	GAGTTTTTTCCAACCCAGGGTTCGTGGCATCTGCATTCGCATCTGTGCCTTGATATTTTGG	2130
Qy	734	IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla	753
Db	2131	GTGGAGACATTATTGTCACATACAGTCTTCCCTGTGATGCTGAATGCTATTGGTCTAGAA	2190
Qy	754	GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys	773
Db	2191	GGCGTGTTTGGCATATACGAGTTGCCCTGTGCCATTGCCTTTGTGTTGTCTATCTTAAG	2250
Qy	774	ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla	793
Db	2251	GTTCCGGAGACAAAAGGCATGCCCTTGAGTGCATCACTGAGTTCTTTGCAGTTGGTGTGCA	2310
Qy	794	LysGlnAlaAlaLalysAla 800	
Db	2311	AAGCAAGCGGTTGCAAAAGGCC 2331	
RESULT 4			
AK099716			
LOCUS			
DEFINITION			
Oryza sativa (japonica cultivar-group) cDNA clone:J013086B19, full insert sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team.; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shihiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group.; Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN.; Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.			
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice			
Science 301 (5631), 376-379 (2003)			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
2 (bases 1 to 2378)			
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kawagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,			

Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niihara, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers  
1. .2378  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J013086B19"

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 3.38e-197 Length: 2378  
Score: 3097.00 Matches: 596  
Percent Similarity: 93.20% Conservative: 34  
Best Local Similarity: 88.17% Mismatches: 46  
Query Match: 76.00% Indels: 0  
DB: 8 Gaps: 0

US-10-051-909-32 (1-800) x AK099716 (1-2378)

Qy	123	AspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGly	142
Db	1	GACTGGATCGCGCGCCCTATGCTCATTTCTCTTCAATTCTTACTTCTCTCAGCAGC	60
Qy	143	LeuIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheValAspGly	162
Db	61	CTCATATGCTATGGTCCCCCGAATGTCTATGTATTACTGCTCGCAGCGCCTTATAGATGGA	120

Qy	163	PheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSer	182
Db	121	TTCGGCATCGGCTTGGCTGTTCACACTTGTACCTTTGTATCATCTCAGAGACAGCTCCTTCA	180
Qy	183	GluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeu	202
Db	181	GAGATCAGGGGTTTGTGTAATACACTGCCACAGTTCACTGGGTCTCAGTGGGATGTTCTTG	240
Qy	203	SerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeu	222
Db	241	TCCTACTGCATGGTGTGGGATGTCACTGTTGCCGTCACTGACTGGAGAATTATGCTT	300
Qy	223	GlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGlu	242
Db	301	GGTGTCTCGCAATCCCTTCGTTGTTTCTTCGGATTGACAAATATTCTATCTGCCAGAA	360
Qy	243	SerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLys	262
Db	361	TCACCAAGATGGCTTGTCAGCAAAAGGGCGGATGGCTGAGGCAAGAAGGTATTACAAAAA	420
Qy	263	LeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluVal	282
Db	421	TTACGTGGGAGAGAGGATGTCTCAGGAGAAATGGCTCTTCTTGTGAAGGTTTGGAGGT	480
Qy	283	GlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAsp	302
Db	481	GGAGCTGACACTTCCATTGAAGAGTACATCATCGGACCGCTATAGAGCCAGCTGATGAG	540
Qy	303	LeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSer	322
Db	541	CATGTTGTTGATGGCGATAAGGACCAATAACACTGTACGGGCTGAAGAGGGCCAATCA	600
Qy	323	TrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer	342
Db	601	TGGATTGCTCGACCTTCCAGGGACCCAGCATTTCTTGGAGTGTGCTTCTCTTACATCT	660
Qy	343	ArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPhe	362
Db	661	CGTCATGGCAGCATGGTGAAACAGAGCGTGCCACTTATGGATCCTATAGTCACGCTTTC	720
Qy	363	GlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhePro	382
Db	721	GGCAGTGTCCATGAGAACATGCCTCATGCTGGAGGAAGCATGCGGAGTACATTGTTTCCC	780
Qy	383	AsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGlu	402
Db	781	AACTTGGTAGTATGTTTAGTGTGACAGATCAGCACCCCAAGGTTGATCAATGGGATGAG	840
Qy	403	GluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGlu	422
Db	841	GAGAACCTTCATAGGGATGATGAGGAATACGCGTCTGTGTTGGTGGAGGTGACTATGAA	900
Qy	423	AspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle	442
Db	901	GACAATGTCCACAGCCCCGTGCTGTCCCAGACAGACCAGCGCAGAAAGGAGACATT	960
Qy	443	ValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGlu	462
Db	961	GCACACCATGCTCATCGTGAAGTGTCTCTGAGTATGAGAAGAAGAGCCTCTTGGAAAGAG	1020
Qy	463	GlyGlyAspGlyValSerSerThrAspIleGlyGlyTyrTrpGlnLeuAlaTrpLysTrp	482
Db	1021	GGTGGCAGGGCGGTGAGCAGCACTGGCATTTGGTGGGGGATGGCAGCTTGCAATGGAATGG	1080
Qy	483	SerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHis	502
Db	1081	TCAGAGCGAGAGGGCGAGGATGGTAAGAAGGAGGAGGATTTAAAGGATCTACTTGCAC	1140
Qy	503	GlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAsp	522
Db	1141	CAAGAGGAAGTCCAGGATCAAGAAGGGGCTCAGTTATTTCACCTTCCTGGTGGAGGGGAC	1200

QY	523	ValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPhe	542
Db	1201	GCTCCTGAAGCAGCGAATTCATACATGCTGCCGCTTTGGTAAGCCAAACCAGCACTTTAC	1260
QY	543	SerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVal	562
Db	1261	TCCAAGGATATCATTTGAACAGCGTATGTCGGTCCAGCCATGATTCATACATCAGAGGCA	1320
QY	563	AlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeu	582
Db	1321	GCTGCCAAAGGTTCAAGCTGGAAGATTGTTTGAACCTGGAGTGAGGCGTGCCTTGTTA	1380
QY	583	ValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyr	602
Db	1381	GTCGGTGTGGAATTCAAATCCTCAACAGTTTGCTGGAATAAATGGGGTTCTCTATTAC	1440
QY	603	ThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSer	622
Db	1441	ACTCCACAAATACTTGAGCAAGCGGGGTGGCAGTTCTCCTTTCCAATCTTGGCCTCAGT	1500
QY	623	SerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle	642
Db	1501	TCAGATCAGCTTCCATCTTGATCAGTTCTCTGACCACCCCTACTGATGCTTCCTAGCATT	1560
QY	643	GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIle	662
Db	1561	GGTTTAGCCATGAGACTTATGGACATCTCTGGAAGAAGGTTTCTGCTTCTGGGCACAATT	1620
QY	663	ProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThr	682
Db	1621	CCAGTCTTGATAGCATCTCTAGTTGTCTTGGTTGTGCCAATGTTATCGACCTGGGTACA	1680
QY	683	LeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMet	702
Db	1681	GTGSCCCACGCCGCACTCTCCACAATCAGCGTCTCATCTACTTCTGCTGCTTCGTCATG	1740
QY	703	GlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGly	722
Db	1741	GGATTCGGTCCGATCCCAACATTCTGTGTGCGGAGATCTTCCCCACTAGGGTCCGCGGC	1800
QY	723	LeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSer	742
Db	1801	ATCTGCATTGCCATCTGCGCCCTGACATTCIGGATTGGTGATATCATAGTCACCTACAGC	1860
QY	743	LeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValVal	762
Db	1861	CTCCCTGTGATGCTGAATGCCATCGGCCTAGCAGGTGTCTTTGGTATATACGCAGTTGTT	1920
QY	763	CysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeu	782
Db	1921	TGCTCGATTGCCTTCGTGTTCTGCTTCCCTCAAGGTCCCCGAGACGAAAGGCATGCCGCTC	1980
QY	783	GluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaAla	798
Db	1981	GAAGTCATCACCGAGTTCTTCGCTGTTGGTGCAAAGCAAATGCAGGCT	2028
RESULT 5			
AP005756			
LOCUS			
DEFINITION			
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,			
BAC clone: OSJNBb0035N08.			
ACCESSION			
AP005756			
VERSION			
AP005756.3 GI:49388934			
KEYWORDS			
Oryza sativa (japonica cultivar-group)			
SOURCE			
Oryza sativa (japonica cultivar-group)			
ORGANISM			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
Erihartioideae; Oryzeae; Oryza.			
REFERENCE			
1			
Sasaki,T., Matsumoto,T. and Katayose,Y.			
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC			
clone:OSJNBb0035N08			

JOURNAL	Published Only in Database (2002)
REFERENCE	2 (bases 1 to 136267)
AUTHORS	Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLE	Direct Submission
JOURNAL	Submitted (18-SEP-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT	On Jun 28, 2004 this sequence version replaced gi:42627749. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark.hmm (http://opal.biology.gatech.edu/GeneMark/), GlimmerM (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), _SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DDBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTp. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from M13rev to -21M13 of the BAC clone. This sequence of OSJNBb0035N08 clone has an overlap with P0620H05 (DDBJ: AP005394) clone at 5' end and with OJ1711.D06 (DDBJ: AP004857) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES	Location/Qualifiers
source	1..136267
	/organism="Oryza sativa (japonica cultivar-group)"
	/mol_type="genomic DNA"
	/cultivar="Nipponbare"
	/db_xref="taxon:39947"
	/chromosome="2"
	/clone="OSJNBb0035N08"
gene	complement(join(157..256,518..608,645..687,1665..1820))
	/gene="OSJNBb0035N08.1"
misc_feature	complement(join(157..256,518..608,645..687,1665..1820))
	/gene="OSJNBb0035N08.1"
	/note="hypothetical ORF predicted by GENSCAN
	this category is not included in IRGSP standard"
gene	complement(6895..7272)
	/gene="OSJNBb0035N08.2"
misc_feature	complement(6895..7272)
	/gene="OSJNBb0035N08.2"
	/note="hypothetical ORF predicted by GlimmerM
	this category is not included in IRGSP standard"
gene	complement(11805..13042)
	/gene="OSJNBb0035N08.3"
mRNA	complement(join(<11805..12148,12855..>13042))
	/gene="OSJNBb0035N08.3"
	/note="supported by full-length cDNA(s): AK111070"
misc_feature	complement(11805..13042)
	/gene="OSJNBb0035N08.3"





QY 81 ----- 81  
Db 48860 GTGTCGTCTAGATTGGTTATTATTGGTTGAGAGTTTCTGTGAGTACTAGGAATATGAACAGC 48919  
QY 81 ----- 81  
Db 48920 TGTCAATTTTCTTAAATATATCAGTGTGGTGATCCTCTCAAATTTTTCAGATTAATAGA 48979  
QY 81 ----- 81  
Db 48980 CTGCTTTAGATCAGATGTTTGTGTGAGTGCTAATTATGGTCTTCTAGTAAACATGTCTGGTG 49039  
QY 81 ----- 81  
Db 49040 TTGTTGTTTGAGTCCTAAGAGCAAATTGGCACAAATTGTTTATCAGCAAACCTGGATCGCATG 49099  
QY 81 ----- 81  
Db 49100 ATATTAATAGTCTCTTAATTTTCTAGATCTGTACTAATCTTATTGAAGAAAGAGATGTT 49159  
QY 81 ----- 81  
Db 49160 CTAATCATAATTGTTAGCTATTGTTGTTGCTGTGTAATCATTACTGTTTGCCATTCTTTG 49219  
QY 81 ----- 81  
Db 49220 TTCTGAACACCTGTAAAGATGTTCTATCCAAAGAAAGTGTAAATGTTAAACAAGGCAC 49279  
QY 82 ----- AlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95  
Db 49280 AANTATTAAACTTCTTCAGGTGCTGTACTATACATAAAGAAGAAATCAAGCTAGAAAGT 49339  
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
Db 49340 GAGCCCACTGTGGAGGGCTAATCGTGGCCATGTCACTGATTGGTGCAACCATCATCACA 49399  
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSerSer 135  
Db 49400 ACATTCTCAGGGCCGGTATCAGACTGGATCGGCGGGCGCCCTAIGCTCATTTCTCTTCA 49459  
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrPheProAsnValTyrValLeuLeu 155  
Db 49460 ATTCTCTACTTCCCTCAGCAGCCTCATCATGTATGTTCCCGAATGCTATGTGTTACTG 49519  
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
Db 49520 CTGCGACGCCTTTATAGATGGATTTCGGCATCGGCTGGCTGTACACTGTACCTTTTGATC 49579  
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
Db 49580 ATCTCAGAGACAGCTCCTTCAGAGATCAGGGGTTTGCTGAATACACTGCCACAGTTTCAGT 49639  
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
Db 49640 GGGTCAGGAGGGATGTTCTTGTCCTACTGCATGGTGTGTTGGGATGTCACGTGTGCCGTCA 49699  
QY 216 ProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhePheGlyLeu 235  
Db 49700 CCTGACTGGAGAAATTATGCTTGGTGTCTTCGCAATCCCTTCGTTGTTTCTTCGGGATG 49759  
QY 236 ThrIlePheTyrLeuProGluSerProArgTyrPheLeuValSerLysGlyArgMetAlaGlu 255  
Db 49760 ACAATATTCTATCTCCAGAAATCACCAAGATGGCTTGTGCAGCAAAGGCGGATGGCTGAG 49819  
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSer----- 270  
Db 49820 GCAAGAAGGTTATTACAAAAAATTACGTGGGAGAGAGGATGTCTCAGGTATGTTTTGCTGA 49879  
QY 270 ----- 270  
Db 49880 AAATTTTGATGTTTCATTTTCACCCATATTTCCAGCTGACCAGTTTAAAAACAATTTCTCCC 49939

QY 270 ----- 270  
Db 49940 AGATTCTAAACTCATGGTTAATACTAGAAAAACATGTGTGGCATTTGGTAACTGTTAAATTC 49999  
QY 270 ----- 270  
Db 50000 TTGACACGAACTTAAACATATGCAGAAATAGTGTATTTTGGAAAAAAGGGTAAAGAAA 50059  
QY 270 ----- 270  
Db 50060 TCAAAAATCTATTAAAGCGCCTTCTATCGGTGCTCTTTTGGAAAAAGCTTATGAATCTTCAAG 50119  
QY 270 ----- 270  
Db 50120 TAAAAAATTACTAAGTACACATAAACTACACTAAATTCATTTTCCATAGAAGGCAGAACATGG 50179  
QY 270 ----- 270  
Db 50180 AATGTACAATCGTTATCTCACCAGGTAGAGCCCCAAATGAGGTGTCCCAACTTTTGTCTGA 50239  
QY 270 ----- 270  
Db 50240 GTCCCTAATCTGGAGAAAAAATAATCAGGTGATCTAGATAGACATCGCCCTCTTTTTTTTT 50299  
QY 270 ----- 270  
Db 50300 TTTTGGTAAGGTAACTCTAGAAATTTGGGGGGTGGGACCTGCTGCTGTATTAATGTGGATGTG 50359  
QY 270 ----- 270  
Db 50360 GGCCGACCCACGGGTCCCTTATCTACTCTAGTTTAACTAAATGAACATATATTTTGCCTTTG 50419  
QY 270 ----- 270  
Db 50420 AGATTTTAAAGTGAATTAAGTCCATCTAGTTTGACCTAGAGTGGGTCAAGGAGCCCCAC 50479  
QY 270 ----- 270  
Db 50480 AGCGGCCATGCTAAACAACAGTGATTGTTTCAGAAACGTTAGTAAATGATTATCCTTAAGT 50539  
QY 270 ----- 270  
Db 50540 TCATGCTTTCCTCAAACTAATTAACAATAACATACATGTGTCAACATCAAAATAGGCATATGCT 50599  
QY 270 ----- 270  
Db 50600 GTTAGCAAGTGGTTTCTCTATTTTCTCTCTACAAACAGACTAAATGGACAGTATTCGGT 50659  
QY 270 ----- 270  
Db 50660 GGCTGAGAGCATAATATTGCTGTATTATTAAAGTTTCATCGAACTTCGCTGTAAGCTGTTG 50719  
QY 270 ----- 270  
Db 50720 TCCCCAGCTGCACATATGCTAGCTTTCTATTAAAGCTGGGCCCTTATATTCTGTATTATG 50779  
QY 270 ----- 270  
Db 50780 TTTCTGGCACACCATTCATATTTTATATTGCTGCTTTATTTCATAGTTACTACTTATTATG 50839  
QY 270 ----- 270  
Db 50840 CCAATGTTTATACCAATGTCCAAATFAAAATCATGTAATCTACGTGAAAAATAAAGTTGC 50899  
QY 271 ----- GlyLeuLeuSerLeuLeuGlyLeuGly 281  
Db 50900 ATAATCAITCAAATTATCAAATTTGGCAGGAGAAATGGCTCTTCTTGTGAAGGTTTGA 50959  
QY 281 uValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAs 301  
Db 50960 GGTGAGCTGACACATTCATTGAAGAGTACATCATCGGACCGGCTATAGAGCCAGCTGA 51019  
QY 301 pAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGlyGly 321

Db 51020 TGAGCATGTTGTTGATGGCGATAAGGACCAATAACACTGTACGGGCCTGAAGAGGGCCA 51079  
Qy 321 nSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAl 341  
Db 51080 ATCATGGATTGCTCGACCTTCCAAAGGACCCAGCATTTCTTGGAGTGTGCTTCTCTTAC 51139  
Qy 341 aSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLe 361  
Db 51140 ATCTCGTCATGGCAGCATGGTGAACAGAGCGTGCCACTTATGGATCCTATAGTCACGCT 51199  
Qy 361 uPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPh 381  
Db 51200 TTTTCGGCAGTGTCCATGAGAACATGCCTCATGCTGGAGGAAGCATGCGGAGTACATTGTT 51259  
Qy 381 eProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAs 401  
Db 51260 TCCCAACTTGGTAGTATGTTTAGTGTGACAGATCAGCACCCCAAGGTTGATCAATGGGA 51319  
Qy 401 pGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTy 421  
Db 51320 TGAGGAGAACCTTCATAGGGATGATGAGGAATACGCGTCTGATGGTGGAGTGACTA 51379  
Qy 421 rGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAs 441  
Db 51380 TGAAGACAATGTCCACAGCCCGTTGCTGTCCCGACAGACCACGAGCGCAGAAAGGA 51439  
Qy 441 pileValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGl 461  
Db 51440 CATTGGACACCATGCTCATCGTGGAAGTGCTCTGAGTATGAGAAGAAGCCTCTTGGGA 51499  
Qy 461 yGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLy 481  
Db 51500 AGAGGTGGCGAGGCGGTGAGCAGCACTGGCATTTGGTGGGGATGGCAGCTTGCAATGGAA 51559  
Qy 481 sTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLe 501  
Db 51560 ATGGTCAGAGCGAGAAAGCGGAGGATGGTAAGAAGGAAGGAGATTAAAGGATCTACTT 51619  
Qy 501 uHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGl 521  
Db 51620 GCACCAAGAGGAAGTGCCAGGATCAAGAAGGGGCTCAGTTATTTCACTTCCTGGTGGAGG 51679  
Qy 521 yAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLe 541  
Db 51680 GGACGCTCCTGAAGGCAGCGAAITTCATACATGCTGCCGCTTGTGTAAGCCAACCAGCACT 51739  
Qy 541 uPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGl 561  
Db 51740 TTACTCCAAGGATATCATTTGAACAGCGTATGTCGGTCCAGCCATGATTCATCCATCAGA 51799  
Qy 561 uValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLe 581  
Db 51800 GGCAGCTGCCAAAGGTTCAAGCTGGAAAGATTGTTTGAACCTGGAGTGAGCGTGCCTT 51859  
Qy 581 uLeuValGlyValGlyIleGlnIleLeuGln----- 591  
Db 51860 GTTAGTCGGTGTGGAAATTCAAATCCTTCAACAGGTAGAGATGATACCAATTTTCATAATT 51919  
Qy 591 ----- 591  
Db 51920 CTTTTTTGTACTGTCCAACATTGACGGTGTTTAAGATTTGTGTACTATTAGAATTCACT 51979  
Qy 592 -----GlnPheAlaGlyIleAsnGlyValLeuTyrTyrT 603  
Db 51980 AATTCCCTTTATGTTCTTTCCCTTGTGCAGTTTGTCTGGAATAAATGGGGTTCTCTATTACA 52039  
Qy 603 hrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerS 623  
Db 52040 CTCCACAAATACTTTGAGCAAGCGGGGGTGGCAGTTTCTCCTTTCCAATCTTGGCCTCAGTT 52099  
Qy 623 erAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleG 643

Db 52100 CAGCATCAGCTTCCATCTTGATCAGTTCTCTGACCACCTACTGATGCTTCCTAGCATTG 52159  
Qy 643 lYPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654  
Db 52160 GTTAGCCATGAGACTTATGGACATCTCTGGAAGAAGGTACTAACTTGTTCCTTGGTTT 52219  
Qy 654 ----- 654  
Db 52220 GCATCATCTTAGTACTACCAAAATATACCAACAAGAAATGCACCCCAAGTTAGTATTAAAT 52279  
Qy 655 -----ArgPheLeuLeuLeuGlyThrIleProIleLeuIleA 667  
Db 52280 ACATCATACATTGTTTCCTTTTTAGGTTTCTGCTTCTGGCACAAATTCAGTCTTGATAG 52339  
Qy 667 laSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaL 687  
Db 52340 CATCTCTAGTTGTCTTGGTTGTGTCCAATGTTATCGACCTGGGTACAGTGGCCACGCCG 52399  
Qy 687 euLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProI 707  
Db 52400 CACTCTCCACAATCAGCGTCATCATCTACTTCTGCTGCTTCGTTCATGGGATTCGGTCCGA 52459  
Qy 707 leProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaI 727  
Db 52460 TCCCCAACATTCTGTGTGCGGAGATCTTCCCCACTAGGTCGCGGCATCTGCATTGCCA 52519  
Qy 727 leCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetL 747  
Db 52520 TCTGCGCCCTGACATTCTGATTGGTGTGATATCATAGTCACCTACAGCCTCCCTGTGATGC 52579  
Qy 747 euAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerP 767  
Db 52580 TGAATGCCATCGGCCTAGCAGGTGTCTTTGGTATATACGAGTTGTTGCTCGATTGCCT 52639  
Qy 767 heValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrG 787  
Db 52640 TCGTGTTCGTCTTCCCTCAAGGTCCCCGAGACGAAGGCGATGCCGCTCGAAGTCATCACCG 52699  
Qy 787 luPhePheAlaValGlyAlaLysGlnAlaAlaA 798

RESULT 6

AK102640

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone:J033100A10, full

insert sequence.

AK102640

VERSION

AK102640.1 GI:32987849

KEYWORDS

FLI\_CDNA; CAP trapper.

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

1

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of  
Agrobiological Sciences Rice Full-Length cDNA Project Team:  
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Hotta,I.,  
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,  
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,  
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International  
Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K.,  
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,  
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,  
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,  
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,  
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:  
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,  
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,  
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,  
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
Yoshino,M. and Hayaahizaki,Y.

AK102640

2800 bp mRNA linear PLN 24-JUL-2003



TITLE  
japonica rice  
JOURNAL Science 301 (5631), 376-379 (2003)  
MEDLINE 22752273  
PUBMED 12869764  
REFERENCE 2 (bases 1 to 2800)  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
This clone is one of the 28K full-length cDNA clones from japonica rice.

COMMENT

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

source

1. .2800  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033100A10"

ORIGIN

Alignment Scores: 5.73e-179 Length: 2800  
Pred. No.: 2825.50 Matches: 560  
Score:

Percent Similarity: 82.32% Conservative: 78  
Best Local Similarity: 72.26% Mismatches: 121  
Query Match: 69.34% Indels: 16  
DB: 8 Gaps: 7

US-10-051-909-32 (1-800) x AK102640 (1-2800)

QY 30 AlaLeuProGlyProLeuProPro-AlaSerCys-----SerSerG1 43  
Db 117 GCCTTCCAGATTCCACGCCGCTCTCTCTTTAGGGATCCGAAATCTCGGTGGACG 176  
QY 43 nGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaValLeuValAl 63  
Db 177 AGAGACTTGGTGGTAAGAT-----TCGCCGGCCATGGCGGCGCGTGTGTCGC 227  
QY 63 aIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAla 83  
Db 228 CATCGCGGCTCCATCGCAACTTGTGCGAGGGCTGGGATATGCAACCATTTGCAGGTGC 287  
QY 83 aValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeu 103  
Db 288 GGTACTGTATCAAGAAGGAATTCACTGCAGAGCGAGCCCTTATCGAAGGCCCTGAT 347  
QY 103 eValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAs 123  
Db 348 CGTGGCCATGTCGTCATTGGGGCGACCATCATCAGACGTTCTCTGGAGCAGTGGCTGA 407  
QY 123 pSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLe 143  
Db 408 TTCTTTTGGTAGCGGCCCATGCTGATCGCGTGGCTGCTCTCTACTTTTGTAGTGGGCT 467  
QY 143 uIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGlyPh 163  
Db 468 AGTGATGCTTTGGGCGCAAAATGTTATGTGTGCTCTTGGCGAGGCTCATTGACGGTT 527  
QY 163 eGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerG1 183  
Db 528 CGGGATCGGTTTGGCTGTACCGTTGTACCATTTGTACATCTCTGAGACTGCCCCCGACGGA 587  
QY 183 uIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSe 203  
Db 588 CATCAGAGGACTGCTAAACACCGTCCGCGCAGTTTCAGTGGGTCTGGAGGGATGTTCTTTC 647  
QY 203 rTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuG1 223  
Db 648 ATACTGCATGGTATTGGCATGTCCCTCATGCCACAGCAGATTTGGAGGATCATGCTTGG 707  
QY 223 yValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSe 243  
Db 708 CGTTCTATCAATACCATCATATATATCTTTGCAATTGACCATCTTTTACTTACCTGAATC 767  
QY 243 rProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLe 263  
Db 768 GCCGAGTGGCTCGTGAGCAAAGGAAGATGGCTGAGGCCAAGCGTGTGTGCAAGGCCT 827  
QY 263 uArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValG1 283  
Db 828 GCGTGGAAAGAGAAGATGTTTCAGGAGAAATGGCCCTTCTCGTTGAAGGTCTGGGGTTGG 887  
QY 283 yGlyAspThrSerIleGluTyrIleIleGlyProAlaThrGluAlaAlaAspLe 303  
Db 888 GAAAGACACAAAATTGAGGAATACATAATTGGACCTGATGATGAGCTTGTGATGAAGG 947  
QY 303 uValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTr 323  
Db 948 GCTGGCTCCA--GATCCAGAGAAGATCAAACCTGTATGGTCTCTGAAGAAGGCTTATCGTG 1004  
QY 323 pIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerAr 343  
Db 1005 GGTGCCCGTCTGTTACGGGCAAGTGCACTTGAAGTGCAATTAGGTCTCATCTCTCG 1064  
QY 343 gHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheG1 363

Db 1065 TCATGGTAGTATGGTCAGTCAGGGTAAGCCCTTGTGGATCCTGTGTCAACCCTTTTGG 1124

Qy 363 ySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAs 383

Db 1125 AAGTGTCCATGAGAAGATGCCTGAGATAATGGGAAGCATGCGGAGCACATGTGTTCCCTAA 1184

Qy 383 nPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluG1 403

Db 1185 CTTTGGCAGCATGTTTAGTGTGGCGGAACAGCAGCAAGCTAAAGGTGATTGGGATGCTGA 1244

Qy 403 uAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAs 423

Db 1245 GAGT---CAACGGGAGGGTGAAGATTATGGATCAGACCATGGTGGGATGACATTGAAGA 1301

Qy 423 pAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle-- 442

Db 1302 TAGCCTCCAAAGCCCACTTATTTCGTCAAGCGACAAGCGTGGAAAGGAAGGAGATCGC 1361

Qy 443 -ValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyG1 462

Db 1362 TGCACCTCATGGCAGTATAATGGGTGCTGTGGGA-----AGAAGTAGTAGTCTCATGCA 1415

Qy 462 uGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTr 482

Db 1416 GGGCGGGAGGCAGTAAGCAGCATGGGCATTGGTGGGGATGGCAGTTGGCTTGGAATG 1475

Qy 482 pSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHi 502

Db 1476 GACTGAGAGAGAAGGTGCAGATGGCGAAAGAAGGTGGCTTCCAACGTATCTACTTGCA 1535

Qy 502 sGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAs 522

Db 1536 TGAAGAGGGTGTGACAGGTGATCGCAGGGCTCTATACTGTCTATTGCCT---GGAGGTGA 1592

Qy 522 pValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPh 542

Db 1593 TGTTCTCTCTGGTGGTGAGTTTCGTCCAGGCAGCTGCTCTTGTCCAGCAACCTGCTCTTTA 1652

Qy 542 eSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVa 562

Db 1653 CTCTAAGGAATTGATGAAGCAACGCCCTTGCTGGCCCTGCTATGGTCATCCATCTCAGGC 1712

Qy 562 lAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLe 582

Db 1713 AGTTGCTAAAGGTCCTCAAAATGGGCAGACTTATTCGAACCTGGAGTGAAGCATGCTCTGT 1772

Qy 582 uValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTy 602

Db 1773 TGTTGGCATAGGGATACAAATCCTGCAACAGTTTGCTGGCATTAAATGGAGTTCTGTACTA 1832

Qy 602 rThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSe 622

Db 1833 CACTCCCAAAATCTTTGAGCAAGCTGGTGTGGTGTCTTCTTGCAAAACATTGGACTTAG 1892

Qy 622 rSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIl 642

Db 1893 CTCCTCATCTGCATCTATTCTTTATTAGCGGACTGACAACCTTGTGACGCTTCCCAGCAT 1952

Qy 642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIl 662

Db 1953 TGGTATTGCTATGAGGCTCATGGATATGTCTGGAAGAAGGTTTCTCTCTCTTGCAACAAT 2012

Qy 662 eProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyTh 682

Db 2013 CCCTATCCTGATAGTACACTAGCTATCTTGATTCTGGTCAATATTCTGGATGTGGGGAC 2072

Qy 682 rLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMe 702

Db 2073 CATGGTTTCATGCCTCACTGTCCACAGTCAGTGTCTACTCTACTTCTGCTTCTTTGTCAT 2132

Qy 702 tGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgG1 722

Db 2133 GGGGTTGGGCGCTATTCCAAACATTTCTGTGCAGAGATTTTCCCGACCAACCGTTCGTGG 2192

Qy 722 yLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSe 742

Db 2193 CATTCGATAGCCATCTGTGCCCTAAACATTCTGGATCGGTGATATCATTTGTGACATAC 2252

Qy 742 rLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValVa 762

Db 2253 CCTCCCGTGATGCTCAACGCCATTGGACTCGTGGAGTGTTCGGAATCTACGCAGTGGT 2312

Qy 762 lCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLe 782

Db 2313 CTGCATACTGGCTTTCCTGTTTGTCTTTCATGAAGGTGCCGGAGACAAAGGGCATGCCTCT 2372

Qy 782 uGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla 796

Db 2373 TGAAGTCATCACCGAGTTCTTCTCTGTCGGAGCAAAAGCAGGCC 2415

RESULT 7

HVU534445 2614 bp mRNA linear PLN 14-FEB-2003

LOCUS Hordeum vulgare mRNA for hexose transporter (stp1 gene).

DEFINITION

ACCESSION AJ534445

VERSION AJ534445.1 GI:26986185

KEYWORDS hexose transporter; stp1 gene.

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.

REFERENCE 1

AUTHORS Weschke,W., Panitz,R., Gubatz,S., Wang,Q., Radchuk,R., Weber,H. and Wobus,U.

TITLE The role of invertases and hexose transporters in controlling ougar ratios in maternal and filial tissues of barley caryopses during early development

JOURNAL Plant J. 33 (2), 395-411 (2003)

MEDLINE 22424051

PUBMED 12535352

REFERENCE 2 (bases 1 to 2614)

AUTHORS Radchuk,R.

TITLE Direct Submission

JOURNAL Submitted (09-DEC-2002) Radchuk R., Molecular Genetics - Gene Expression, Plant Genetics and Crop Plant Research, Corrensstr. 3, Gatersleben, D-06466, GERMANY

FEATURES

Location/Qualifiers

1. .2614

/organism="Hordeum vulgare subsp. vulgare"

/mol\_type="mRNA"

/cultivar="barke"

/sub\_species="vulgare"

/db\_xref="taxon:112509"

/tissue\_type="developing caryopsis"

/clone\_lib="lambda ZAP2"

/dev\_stage="0-15 days after flowering"

1. .2614

/gene="stp1"

83. .2314

/gene="stp1"

/codon\_start=1

/product="hexose transporter"

/protein\_id="CAD58958.1"

/db\_xref="GI:26986186"

/db\_xref="GOA:Q8GT52"

/db\_xref="UniProt/TrEMBL:Q8GT52"

/translation="MSSGAVLVVAIAASIGNLLQGDWNTIAGAVLYIKKFSLETOPL IEGLIVAMSLIGATVITTFSGAVADAVGRRPLLIASSVLYFVSGLVMLWAPNVVYLLL ARLIDFGIGLAVTLPLYISETAPTDIRGLNTLPQFSGSGGMFLSYCMVFTMSLMP QPDWRIMLGVLSPSLMYFALTVFYLPESPRWLVSGRMAEAKRVLQRLRGREDVSGE MALLVEGLGVGKDTHFEEYIIGPDDELADGDLAQEKLKLYGAEEGVSWIARPVRXG QQSALGSALGLMSRHGSMVSGKSLVDPLVTLFGSVHEKMPVGMGSMRSTLPNPFSGM FSVAEQQQAKADWDAESHRDDEYASDHGADDIEDNLNPLISRQATSVEGKEIAAPH GSIIMGVESSMQGDVAVSSMIGGGWQLAWKWTREGADGQKEGGFQRIYLHEEGVS GDRRGSIILMPGGDIPPGGEYIQAALVSQPALYSKDLIEQQLAGPAMVHPSEAVAKG

gene

CDS

TKWAELFEPGVKHALFVIGILOIQFAGINGVLVYTPQILEOAGVGILLSNIGLSSS  
SASILISALTTLMLPSIGIAMRLMDMSGRFFLLSTIPVLIVALAILVLNVLDVGT  
MVHAALSTISVIVYFCFFVMGFGPIPNILCAEIFTSVRGICIAICALTFWIGDIIVT  
YTLPLVMLNAIGLAGVFGIYAVVCMIAFVFVYMKVPETKGMPLVEITEFVSQAKQKE  
ATD"

ORIGIN

Alignment Scores:  
Pred. No.: 7.65e-178 Length: 2614  
Score: 2808.00 Matches: 543  
Percent Similarity: 84.39% Conservative: 84  
Best Local Similarity: 73.08% Mismatches: 104  
Query Match: 68.91% Indels: 12  
DB: 8 Gaps: 6

US-10-051-909-32 (1-800) x HVU534445 (1-2614)

QY	57	SerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAsp	76
Db	89	TCGGGGCCGCTGCTGCCCATAGCGCGCTCCATCGGGAACCTGCTGCAGGGTGGGAT	148
QY	77	AsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGlu	96
Db	149	AATGCGACCATGCGGTGCCGTGTACATAAAGAAGAGATTGAGCCTGGAGACCCAG	208
QY	97	ProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	116
Db	209	CCCTTGATCAGGGCCTCATCGTGGCCATGCTCGCTCATCGGAGCGACGGTTATCAGCAGC	268
QY	117	PheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIle	136
Db	269	TTCTCCGGGCGGTGCTGACGTGTGGTAGCGGCCCTGCTGATCGCTCGTCTGTC	328
QY	137	LeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu	156
Db	329	CTCTACTTGTGAGTGGCCTGGTGATGCTCTGGGCGGCCAACGCTCTATGTGTGCTCTTG	388
QY	157	AlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIle	176
Db	389	GCCAGGCTCATCGACGGGTTCGGTATCGGTTTGGCAGTCACCCCTGTGTCCTGTACATT	448
QY	177	SerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGly	196
Db	449	TCGGAGACCGCCCCGACTGACATTAGAGGGCTGCTGAAACACGCTGCCGCGAGTTCAGTGGG	508
QY	197	SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerPro	216
Db	509	TCAGGAGGATGTTCCCTTCTTACTGTCATGGTGTTACCATGTCCCTCATGCCGCGCCT	568
QY	217	AspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThr	236
Db	569	GACTGGAGAATCATGCTTGGGGT-TTGTGCGATCCCGTCGCTTATGTATTTGCATTGACT	628
QY	237	IlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAla	256
Db	629	GTCCTTTATTGCCCAGTTCGCCCAGATGGCTTGTGAGCAAGGGAAGAAATGGCTGAGGCC	688
QY	257	LysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeu	276
Db	689	AAGCGAGTATTGCAAGACTGCGGGGAAGGAAGATGTCTCAGGAGAGATGGCCCTTCTT	748
QY	277	LeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAla	296
Db	749	GTTGAAGGATTGGGTGTGGAAAAGACACACATTTTGAGGAATACATAATTGGACCTGAT	808
QY	297	ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGly	316
Db	809	GATGAGCTGGCTGATGACGGCTGGCTCCA--GATCAAGAGAAGTTGAAACTCTATGGA	865
QY	317	ProGluGluGlyGlnSerTrpIleAlaArgProSerLys-----GlyProIleMetLeu	334
Db	866	GCTGAAGAAGGGGTATCTTGGATCGCCCGTCCCTGTTAGGSGTGGCGGCCAAAGTGCACTT	925

QY	335	GlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeu	354
Db	926	GGAAGCGCTTTGGCCCTCATGCTCGTTCATGGAGTATGGTTAGTCAGGGTAAATCTCTT	985
QY	355	MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly	374
Db	986	GTTGACCCCTTGTCACTCTCTCGGAAGTGTCCATGAGAAGATGCCTGAGGTAATGGGG	1045
QY	375	SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis	394
Db	1046	AGCATGCGGAGCACCTTGTTCCTAACTTTGGCAGCATGTTTAGTGTGGCTGAACAGCAG	1105
QY	395	AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer	414
Db	1106	CAGGCTAAAGCTGACTGGGATGCTGAGAGT--CATAGGGATGATGAAGATTATGCATCG	1162
QY	415	AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla	434
Db	1163	GATCATGGTGTGATGACATTGAGGATAACCTCAATAGCCCGCTATTCTCGTCAAGCG	1222
QY	435	ThrGlyAlaGluGlyLysAspIleVal-----HisHisGlyHisArgGlySerAlaLeu	452
Db	1223	ACAAAGCGTGGAGGCTAAGGAGATTGCTGCACCTCATTGGAAAGCATAATGGGTGGTGGAA	1282
QY	453	SerMetArgArgGlnThrLeuLeuGlyGluGlyValSerSerThrAspIle	472
Db	1283	AGTAGTAGCATGCAG-----GGAGGGGACGCAAGCAGCATGGGCATT	1327
QY	473	GlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLys	492
Db	1328	GGTGGGGGTGGCAGTTAGCTTGGAAAGTGGACTGAGAGAGAAGGTGCAGATGGGCAAAA	1387
QY	493	GluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGly	512
Db	1388	GAAGGTGGCTTCCAGCGTATTACTTGCATGAGAGGGTGTGTCAAGGTATCGGAGAGGC	1447
QY	513	SerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAla	532
Db	1448	TCTATATTGTCTATGCCA--GGAGGTGATATTCTCTCTGTTGGTGGTATATCCAGGCA	1504
QY	533	AlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSer	552
Db	1505	GCCGCTCTAGTGAGCCAACCTGCTCTTACTCGAAGGACCTGATAGAGCAGCAGCTTGCT	1564
QY	553	AspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeu	572
Db	1565	GGTCCAGCCATGGTACATCCATCCGAGGCAGTTGCCAAGGGTACAAAGTGGCGAGAACTA	1624
QY	573	PheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGln	592
Db	1625	TTTGAACCTGGAGTGAAGCATGCACCTGTTTGTGGCATTTGGATTACAGATCCTGCAGCAG	1684
QY	593	PheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyVal	612
Db	1685	TTTGGGGTATCAACGGAGTCTGTATTACACGCCCTCAGATACTTGAGCAAGCAGGTGTT	1744
QY	613	AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSer	632
Db	1745	GGGATTCTTCTATCAAAACATTGGACTAAGCTCTTCTCTGCATCTATTCTTATTAGTGCC	1804
QY	633	LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer	652
Db	1805	TTGACAACCTTGCTGATGCTTCCCAGCATTTGGCATTCGCCATGAGACTCATGGATATGTCA	1864
QY	653	GlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu	672
Db	1865	GGAAGAAGGTTTCTTCTCTCTCAACAATCCCTGTCTTGTATAGTGGCCTAGCTATCTTG	1924
QY	673	ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer	692
Db	1925	GTTTGGTGAATGTTCTGGATGTGGGAACCATGGGTGCACGCCGCTCTCTCAACGATCAGT	1984
QY	693	ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys	712



|||||  
Db 1985 GTCATAGTCTACTTCTGCTTCTTCGTCATGGGTTTGGACCTATCCCGAATATTCTCTGC 2044  
713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732  
|||||  
Db 2045 GCGGAGATTTCGCCACCTCCGTCGTCATCTGCATTGCCATCTGCGCACCTAACCTTC 2104  
733 TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752  
|||||  
Db 2105 TGGATTGGTGACATCATCGTGACATACACTCTTCCCGTGATGTCTCAACGCCATCGGACTC 2164  
753 AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772  
|||||  
Db 2165 GCTGGAGTTTTCGGAATATATGCTGCTGTTGCATGATAGCCTTCGTATTCGTCTACATG 2224  
773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792  
|||||  
Db 2225 AAGGTCCTGAGACAAAGGGCATGCCCTGGAGGTCATCACCGAGTCTTCTCGGTCCGG 2284  
793 AlaLysGln 795  
|||||  
Db 2285 GCGAAGCAG 2293

RESULT 8  
AR208568  
LOCUS AR208568 2601 bp DNA linear PAT 20-JUN-2002  
DEFINITION Sequence 7 from patent US 6383776.  
ACCESSION AR208568  
VERSION AR208568.1 GI:21509755  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2601)  
AUTHORS Allen,S.M., Hitz,W.D., Kinney,A.J. and Tingey,S.V.  
TITLE Plant sugar transport proteins  
JOURNAL Patent: US 6383776-A 7 07-MAY-2002;  
FEATURES  
source  
1. .2601  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 6.72e-169 Length: 2601  
Score: 2674.00 Matches: 522  
Percent Similarity: 82.20% Conservative: 92  
Best Local Similarity: 69.88% Mismatches: 115  
Query Match: 65.62% Indels: 18  
DB: 6 Gaps: 8

US-10-051-909-32 (1-800) x AR208568 (1-2601)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75  
|||  
Db 175 ATGAAAGTGCCGTCCTTGTGCTATTGCCGCTTCCATTGGTAATTTCCTCCAAGGATGG 234  
76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95  
|||  
Db 235 GATAATGCTACCATTGCCGGGGCTAATGTTACATTAAAGAAAGACCTTGCTTTGGGAACA 294  
96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
|||  
Db 295 -----ACTATGGAAGGCTTGTGGTGCGCATGTCCCTGATTGGAGCAACGGTAATCACC 348  
116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
|||  
Db 349 ACATGCTCTGGTCCTATAGCGGATTGGCTCGGTCGGCGACCCATGATGATAATCTCATCT 408  
136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155  
:::  
Db 409 GTGCTCTATTCTTGGGTGGTTTGGTGATGCTGTGGTCCCCCAATGTGTATGTGTGTGTC 468

Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
|||||  
Db 469 TTGGCGAGGCTACTTGTATGGATTGGGATTGGCCTTGCCTGACTCTTGTCCCGTCTAT 528  
176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
|||||  
Db 529 ATATCTGAACGGCGCGCTCTGAATAAGGGGTCGTTTGAATACGCTTCTTCAGTTCAGT 588  
196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
|||||  
Db 589 GGCTCTGGAGGAATGTTTTTGTCTGCTACTGTATGGTTTTTGGCATGTCTAGTCCCGCG 648  
216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
|||  
Db 649 CCTAGCTGGAGGCTCATGCTTGGGGTTCTGTCTATTCTCTCTCTCTTTTGCATTG 708  
236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
|||||  
Db 709 ACCATTTTTTCTTGCCCGAGTCTCTCGGTGGCTGGTCAGCAAGGAAGGATGCTCGAG 768  
256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
|||||  
Db 769 GCTAAGAAGGTGCTCCAAGATTGCGGGAAGGAGGATGTGTCTCAGCGGAGATGGCATTG 828  
276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
|||  
Db 829 CTGGTTGAAGGTCTCGGGATTGGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT 888  
296 AlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
|||  
Db 889 GCTGACGATGTGGCTGATGGTGCATGAACATGCAACAGAGAAGATAAAATTCGATTATAT 948  
316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
|||  
Db 949 GGATCCCAAGCAGGCCCTTCTTGGTTATCAAAACCTGTCACTGGACAGAGTTCTTATTGGC 1008  
336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355  
|||||  
Db 1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056  
356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373  
|||||  
Db 1057 GATCCTCTGGTGACACTGTTGGTAGCATTTCATGAGAAGCTCCCCGAGACAGGCAAGA 1116  
374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
|||||  
Db 1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGAAGCATGTTTCAGCACTGCTGAGCCG 1176  
394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413  
|||||  
Db 1177 CATGCTAAAAATTGAACAATGGGATGAAGAAAGCTTACAAGGGGAACGTGAGGACTACATG 1236  
414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433  
|||||  
Db 1237 TCAGATGCAACCCGTGGGGACTCCGATGATAATTGCAAGTCTTAAATCTCACGCCAA 1296  
434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452  
|||  
Db 1297 ACAACAAGCCTTGAA--AAAGACTTACCTCCTCCTCCTTCCCATGGCAGTATCCTTGGC 1353  
453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471  
|||||  
Db 1354 AGCATGAGCGCTCACAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413  
472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGluAsnGlyArg 491  
|||||  
Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAAATGGACTGATAAA--GGTGAGGATGGAAAA 1470  
492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511  
:::  
Db 1471 CAACAAGGAGGGTTTAAAGGATTATTACATGAGGAGGAGGTTTCTGCATCTCGTCGT 1530  
512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531

Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGC-----GAATTGTCCAG 1572

Qy 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551

Db 1573 GCTGCTGCCTTGGTAAAGCCAACCCGCTCTTTACTCCAGGAGCTTATTGATGGACACCCA 1632

Qy 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571

Db 1633 GTTGGGCCTGCAATGTTCACCCATCTGAGACAGCTTCAAAGGGGCCAAGTTGGAAGCT 1692

Qy 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591

Db 1693 CTTCTTGAACCAAGGGTTAAGCATGTCATTTGGTTGTTGAGTTGGAATACAATACTTCAG 1752

Qy 592 GlnPheAlaGlyIleAsnGlyValLeuTyrThrProGlnIleLeuGluGlnAlaGly 611

Db 1753 CAGTTTTCAGGGATAAATGGGTTCTATATTACACACTCAAATCCTTGAAGAGGCCGGT 1812

Qy 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631

Db 1813 GTTGAAGTTCTTCTTCAGATATAGGCATTTGGCTCAGAGTCGGCATCATCTCTTATCAGT 1872

Qy 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651

Db 1873 GCITTCACAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGGATGTT 1932

Qy 652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671

Db 1933 TCAGGCAGAAAGGCAGTTGCTACTTACTACAATCCCGTGTGCTGATTGTGTCACTCATTT 1992

Qy 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691

Db 1993 TTGCTCATTTGGAAGCCCTGGTAAATTTTGGCAATGTCGCCCATGCGCAATCTCAACAGTA 2052

Qy 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711

Db 2053 TCGCTGTGTTTATTCTGCTGCTTTGTGATGGGTTATGGACCAATCCAAACATCCTT 2112

Qy 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731

Db 2113 TGCTCAGAGATTTCCCCACTAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2172

Qy 732 PheTrpIleGlyAspIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751

Db 2173 TTCTGGATTGGAGACATCATCATCATCTACATCTCGCTGCCTGTGATGCTCGGCTCTTTAGGA 2232

Qy 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771

Db 2233 CTTGGTGGTGTATTTCGCCATTACGCAGTTGTTTGTTCATCTCTCGTGATATTTGTGTT 2292

Qy 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791

Db 2293 TTGAAGGTTCCAGAAACAAAGGGCATGCCCTTGAAGTCACTCTCTGAAATTCCTTTCTGTT 2352

Qy 792 GlyAlaLysGlnAlaAlaAla 798

Db 2353 GGAGCAAGCAGGCTGCTTCT 2373

RESULT 9

AY094465

LOCUS

DEFINITION Arabidopsis thaliana AT4g35300/F23E12\_140 mRNA, linear PLN 05-MAY-2002

ACCESSION AY094465

VERSION AY094465.1 GI:20453188

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 2570)

AUTHORS Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J.,

Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

Arabidopsis cDNA clones

Unpublished

2 (bases 1 to 2570)

Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

Direct Submission

Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory (SIGNAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEN (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shinn,P., Chen,H., Cheuk,R., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Chang,E., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura, Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Shinn,P. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers

1..2570

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="4"

/clone="RAFL09-48-J19 (R21968)"

/ecotype="Columbia"

209..2428

/note="putative sugar transporter protein"

/codon\_start=1

/product="AT4g35300/F23E12\_140"

/protein\_id="AAM19835.1"

/db\_xref="GI:20453189"

/translation="MSGAVLVVAIAAAVGNLLQGDNATIAGAVLYIKKEFNLESNPSV  
EGLIVAMSLIGATLITTCGGVADWLGRRPMLILSSILYFVGSVLMLWSPNVYVLLG  
RLLDGFGVGLVTVPIYISETAPPEIRGLLNTLPQFTGGMFLSYCMVFGMSLMPS  
PSWRLMLGLFIPSLVFFFLTVFFLPESPRWLVSKGRMLEAKRVLQRLRGREDVSGEM  
ALLVEGLGIGGETIEEYIIGPADEVTDHDIADKQIKLYGAEGLSWVARPVKGG  
STMSVLSRHGSTMRRQGSLLIDPLVTLFGSVHEKMPDTGSMRSALFPHFSGMFSVGGN  
QPRHEDWDEENLVGEGEDYPSDHGDDSEDDLSPLISRQTTSMKDMPTAHGTLSTF  
RHGSQVQAGGEGAGSMGIGGGWQVAKWTEREDSGQEGGFKRIYLHQEGFPGSRR  
GSI VSLPGSGDGTGADFVOASALVSQALYSKDLLKEHTTIGPAMVHPSETTKGSIWHD  
LHDPGVKRALVVGVLQILQQFSGNGVLYYTPQILEQAGVIGLLSNMGISSSASLL  
ISALTTFVMLPAIAVAMRLMDLSGRRTLLTPIPIASLLVVISNLVHMNSIVHAV  
LSTSVSVLYFCFFVMGFGPAPNILCSEIFPTRVRGICIAICALTFCWICDIIIVTSLPV  
LLKSI GLAGVFGMYAIVCCISWVFVFIKVPETKGMPLFVTEFFSVGARQAEAAKNE"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

9.01e-168

2657.00

Length:

2570

Matches:

530

Percent Similarity: 79.87% Conservative: 97  
Best Local Similarity: 67.52% Mismatches: 133  
Query Match: 65.20% Indels: 25  
DB: 8 Gaps: 10

US-10-051-909-32 (1-800) x AY094465 (1-2570)

Qy	29	LeuAlaLeuProGlyProLeuProAlaSerCysSerSerGlnGlu-----	44
Db	91	CTTCTTCTGCCT---CCAATTCCAAGTGGATTTGTTCACCTACCTCATCGTCGTTGATC	147
Qy	45	-----ProValThrSerAspAspIleLeuGluAspLys-----	55
Db	148	AGTGACTCTGTTTTTGGAAATTCCTCTCAGATTCTTGATAAAAGATAGAACGGAAT	207
Qy	56	-MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTr	75
Db	208	TATGAGTGGAGCTGTGCTTGTGTGCTATTGCTGTGCTGTGTTGGCAACTGTTACAAGGATG	267
Qy	75	pAspAsnAlaThrIleAlaAlaValIleuTyrIleLysLysGluPheGlnLeuGlnAs	95
Db	268	GGATAACGCAACTATTGCAGGAGCTGTGTGTACATAAAAAAGGAGTTTAATTTGGAGAG	327
Qy	95	nGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValTh	115
Db	328	TAATCCATCAGTGAAGGTCTAATTGTGCGGATGTCACTTATTGSGTACTCTGTATTAC	387
Qy	115	rThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSe	135
Db	388	AACATGCTCTGGAGGGGTAGCTGATTGGCTTGGTCGCCGTCCCATGCTAATATTGTCTC	447
Qy	135	rIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLe	155
Db	448	AATTCTCTACTTTGTGGTTCTCTAGTAATGCTATGGTCTCCGAATGTTTATGTGTGCT	507
Qy	155	uLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTy	175
Db	508	CTTAGGAAGGTTGTAGATGGATTTGGGTTGGTCTTGTGTGCACACTTGTTCCTATTTA	567
Qy	175	rIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSe	195
Db	568	TATATCTGAGACTGCACCACCTGAGATTAGGGGACTGTTGAATACGTTACCGCAGTTTCA	627
Qy	195	rGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSe	215
Db	628	TGGCTCTGGAGGGATGTTCTTATCTTACTGTATGGTTTTCGGAATGTCGTTGATGCCATC	687
Qy	215	rProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLe	235
Db	688	ACCTAGCTGGAGATTGATGCTTGGTGTCCTTTTCATCCCTTGCCTTTTCTTCTTCCCT	747
Qy	235	uThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGl	255
Db	748	CACGGTCTTCTTCTTGTCCCGAGTCCCAAGTGGCTCGTGAGCAAAAGGTCGAATGCTTGA	807
Qy	255	uAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLe	275
Db	808	AGCAAAGCGGGTTCTTCAGAGACTGCGTGTGCGGAAGATGTGTCTGCTGAGATGGCTTT	867
Qy	275	uLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPr	295
Db	868	GTTGGTTGAGGGTCTTGGAAATTGGAGGTGAACAACCATAGAGGAATATATAATTGGTCC	927
Qy	295	oAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTy	315
Db	928	CGCGGATGAAGTTACTGATGATCATGATATAGCTGTGGATAAGGATCAAAATTAAAGTTATA	987
Qy	315	rGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGl	335
Db	988	TGGTGCAGAAAGGGCTGAGTTGGGTTGCTAGGCCAGTCAAAAGGA-----GG	1035
Qy	335	ySerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLe	354

Db	1036	AAGCACTATGAGTGTTTTGTCTGCCCATGGAAGTACAATGAGCAGGAGGCAAGGCTCAT	1095
Qy	354	uMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGl	374
Db	1096	GATTGATCCTCTTGTTCACACTGTTTGGGAGCGTTCACGAGAAGATGCCGGACACT---	1152
Qy	374	ySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHi	394
Db	1153	AAGCATGAGGAGTGCCTTGTTCCACATTTTGGGAGTATGTTAGTGTGGAGGAATCA	1212
Qy	394	sAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSe	414
Db	1213	ACCAAGACATGAAGATTGGGATGAAGAGAAATCTTGTGGAGAGGTGAGGATTATCCATC	1272
Qy	414	rAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisserProLeuLeuSerArgGlnAl	434
Db	1273	CGAC---CATGGAGATGATTCTGAAGATGATCTTTCATCTCCGTTGATCTCACGTCAAAC	1329
Qy	434	aThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMe	454
Db	1330	GACAAACATGGAG--AAAGACATGCCCTCACACTGCTCATGGAACTCTTCTACCTTCAG	1386
Qy	454	tArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValserSerThrAspIleGlyGl	474
Db	1387	ACATGGAAGTCAAGTGCAGGGAGCTCAAGGGGAAGGACGGGTAGTATGGGATTGGAGG	1446
Qy	474	yGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGl	494
Db	1447	TGGATGGCAAGTGGCATGGAATGACGCGGAAGAGAGATGAATCGGGACAGAAAGAGG	1506
Qy	494	yGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIl	514
Db	1507	TGGGTTTAAACGGATATACTTGCATCAAGAAGGTTTCCCAAGGATCTCGACGTGGCTCAAT	1566
Qy	514	eValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAl	534
Db	1567	TGTTTCATTCCTGGTGGTGTGATGGAACCGGTGAG--GCAGATTTTGTACAAGCGTCTGC	1623
Qy	534	aLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAl	554
Db	1624	TTTGGTTAGCCAAACCAGCTCTTATTCCAAGACCTTCTCAAAGAACATACAATTGGTCC	1683
Qy	554	aAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGl	574
Db	1684	TGCTATGGTACATCCATCCGAA--ACAACATAAAGGTCAAATTGGCATGATCTTCATGA	1740
Qy	574	uProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAl	594
Db	1741	TCCTGGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAAATACATTTCAGCAGTTCTC	1800
Qy	594	aGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVa	614
Db	1801	AGGCATCAACGGAGTTCCTTTACTACACACCGCAAATCCTTTGAGCAGCGGGTGTCCGGAT	1860
Qy	614	lIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuTh	634
Db	1861	CCTACTATCGAAACATGGGGATTAGTTCTTCTCAGCATCCTTACTTATAAGTGCATTGAC	1920
Qy	634	rThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyAr	654
Db	1921	AACCTTGTGATGTTACCTGCAATAGCTGTTGCAATGAGGCTCATGGATCTTTCTGGTCG	1980
Qy	654	gArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVa	674
Db	1981	AAGGACCTTGCTTCTCACCACGATACCAAATCCTGTATAGCATCTCTATTGTTTAGTAAT	2040
Qy	674	lSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIl	694
Db	2041	CTCAAAATCTTGTTCACATGAACACAGCATGTGCACGCGGTCTTATCAACCGTAAGCGTTGT	2100
Qy	694	eValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGl	714
Db	2101	GCTCTACTTCTGTCTTCTTCGTGTATGGGTTTCCGTCCTGCTCCCAACATCCTCTGTTCAGA	2160



QY	714	uilePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpI	734
Db	2161	GATTTTCCAACTCGAGTCGCGGAATCTGCATCGCATCTGCGCACTCACCTTCTGGAT	2220
QY	734	eGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaG	754
Db	2221	CTGTGACATAATCGTCACTTACAGTCTCCCCGIGCTGCTCAAATCCATTGGACTAGCTGG	2280
QY	754	yValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVa	774
Db	2281	TGTGTTTGGAAATGACGCAATCGTATGTTGCAATTCATGGGTCTTTGTGTTCAATAAGT	2340
QY	774	lProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValIcIyAlaLy	794
Db	2341	CCCCGAAACTAAAGGCATGCCACTTGAAGTCATCACAGAGTTCTTTTCTGTTGGAGCTAG	2400
QY	794	sGlnAlaAlaAla 798	
Db	2401	ACAAGCTGAAGCT 2413	
RESULT 10			
LOCUS	AX506620	2190 bp	DNA linear PAT 27-SEP-2002
DEFINITION	Sequence 1315 from Patent WO0216655.		
ACCESSION	AX506620		
VERSION	AX506620.1	GI:23387857	
KEYWORDS			
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	Harper, J.F., Kreps, J., Wang, X. and Zhu, T.		
TITLE	Stress-regulated genes of plants, transgenic plants containing same, and methods of use		
JOURNAL	patent: WO 0216655-A 1315 28-FEB-2002; The Scripps Research Institute (US); Syngenta Participations AG (CH)		
FEATURES	Location/Qualifiers		
source	1. .2190		
	/organism="Arabidopsis thaliana"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:3702"		
ORIGIN			
Alignment Scores:			
Pred. No.:	5.34e-163	Length:	2190
Score:	2584.00	Matches:	512
Percent Similarity:	80.78%	Conservative:	89
Best Local Similarity:	68.82%	Mismatches:	123
Query Match:	63.41%	Indels:	20
DB:	6	Gaps:	8
US-10-051-909-32 (1-800) x AX506620 (1-2190)			
QY	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75
Db	1	ATGAGTGGAGTGCTGCTTGTGCTATTGCTGCTGCTGTTGGCAACTTGTGTACAAGGATGG	60
QY	76	AspAsnLaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn	95
Db	61	GATAACGCAACTATTGCAGGAGCTGTGTTGTACATAAAAAGGAGTTTAATTGGAGAGT	120
QY	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Db	121	AATCCATCAGTGAAGGTCTAATTGTGGCGATGTCACCTATTGGTGTCTCTGATTACA	180
QY	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgProMetLeuIleLeuSerSer	135
Db	181	ACATGCTCTGGAGGGTAGTGTGATTTGGTTGGTCGCCGCTCCATGCTAATATGTCCTCA	240



Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 301 TTAGGAAGGTTGTTAGATGGAATTGGGGTTGGTCTTGTGGTCACACTTGTCTCTATTAT 360  
  
Qy 176 IleSerGluIleAlaProSerGluIleAArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 361 ATATCTGAGACTGCACCACTGAGATTAGGGGACTGTTGAATACGCTACCGCAGTTCACT 420  
  
Qy 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 421 GGCTCTGGAGGGATGTTCTTATCTTACTGTATGGTTCGGAATGTCGTGATGCCATCA 480  
  
Qy 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 481 CCTAGCTGGAGATTGATGCTTGGTGCTCTTTTCATCCCTTCCCTTGTCTTTCTTCCTC 540  
  
Qy 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 541 ACGGCTCTTCTTGTCCCGAGTCCCAAGGTGGCTCGTGAGCAAGGTCGAATGCTTGAA 600  
  
Qy 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 601 GCAAGCGGGTTCTTCAGAGACTGCGTGGTTCGCGAAGATGTGTCGTGAGATGGCTTTG 660  
  
Qy 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 661 TTGGTTGAGGGTCTTGGAAATTGGAGGTGAACAACCATAGAGGAATATATAATTGGTCCC 720  
  
Qy 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 721 GCGGATGAAGTTACTGATGATCATGATATAGCTGTGGATAAGGATCAAAATTAAGTTATAT 780  
  
Qy 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 781 GGTGCAGAAGAAGGGCTGAGTTGGTTGCTAGGCCAGTCAAAGGA-----GGA 828  
  
Qy 336 SerValLeuSerLeuAlaSerArgHisGlySer--MetValAsnGlnSerValProLeu 354  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 829 AGCACTATGAGTGTGTTGCTCGCCATGGAAGTACAAATGAGCAGGAGCAAGGCTCATTTG 888  
  
Qy 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 889 ATTGATCCTCTTGTTCACACTGTTTGGGAGCGTTACGAGAAGATGCCGGACACT---GGA 945  
  
Qy 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 946 AGCATGAGGAGTGCCTTGTTCCACATTTTGGGAGTATGTTTCAGTGTGGAGGAATCAA 1005  
  
Qy 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1006 CCAAGACATGAAGATTGGGATGAAGAGAAATCTTGTGGAGAAGGTGAGGATTATCCATCC 1065  
  
Qy 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1066 GAC---CATGGAGATGATCTGAAGATGATCTTCATCTCCGTTGATCTCAGTCAAAACG 1122  
  
Qy 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1123 ACAAGCATGGAG---AAAGACATGCCCTCACACTGCTCATGTGAACCTCTTCTACCTTCAGA 1179  
  
Qy 455 ArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly 474  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1180 CATGGAAGTCAAGTCAGGAGGCTCAAGGGGAAGGAGCGGGTAGTATGGGGATTGGAGGT 1239  
  
Qy 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1240 GGATGGCAAGTGGCATGGAATGGACGGAAGAGAAGATGAATCGGACACAGAAAGAA--- 1296  
  
Qy 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1297 -----GAGGTTTCCAGGATCTCGACGTGGCTCAATT 1329

Qy 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAla 534  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1330 GTTTCATTGCCTGGTGGTGATGGAACCGGTGAG---GCAGATTTTGTACAAGCGTCTGCT 1386  
  
Qy 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1387 TTGGTTAGCCCAACAGCTCTTTATTCCAAAGACCTTCTCAAAGAACATACAATTTGGTCCT 1446  
  
Qy 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1447 GCTATGGTACATCCATCCGAA--ACAACATAAAGGGTCAATTTGGCATGATCTTTCATGAT 1503  
  
Qy 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla 594  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1504 CTTGGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAAATACTTCAGCAGTTCTCA 1563  
  
Qy 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1564 GGCATCAACGGAGTTCTTTACTACACACCGCAAATCCTTGAGCAGCGGGTGTCCGGGATC 1623  
  
Qy 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1624 CTACTATCGAACATGGGATTAGTTCTTCTCCTCAGCATCCTTACTTATAAGTGCATTGACA 1683  
  
Qy 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1684 ACCTTTGTGATGTACCTGCAATAGCTGTGTCAATGAGGCTCATGGATCTTTCTGGTCGA 1743  
  
Qy 655 ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1744 AGGACCTTGTCTTCACCACGATACCAATCCTGATAGCATCTCTATTGGTTTTAGTAATC 1803  
  
Qy 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1804 TCAAATCTTGTTCACATGAACAGCATTTGTGCACGCGGTCTTATCAACCGTAAGCGTTGTG 1863  
  
Qy 695 ValTyrPheCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1864 CTCTACTTCTGCTTCTTCGTGATGGGTTTCGGTCTCTGCTCCAAACATCCTCTGTTTCAGAG 1923  
  
Qy 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1924 ATTTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCATCTGCGCACTCACCTTCTGGATC 1983  
  
Qy 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 1984 TGTGACATAATCGTCACCTACAGTCTCCCGGTGCTGTCAAATCCATTTGGACTAGCTGGT 2043  
  
Qy 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 2044 GTGTTTGGAAATGACGCAATCGTATGTTCATTTTCATGGGTCTTTGTGTTCAATAAGTC 2103  
  
Qy 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 2104 CCGGAAACTAAAGGCATGCCACTTGAAGTCATCACAGAGTTCTTTCTGTTGGAGCTAGA 2163  
  
Qy 795 GlnAlaAlaAla 798  
| | | | | : : : | | | | | : : | | | | | | | | | : : | | |  
Db 2164 CAAGCTGAAGCT 2175  
  
RESULT 12  
AK065191  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) cDNA clone:J013002E10, full  
insert sequence.  
AK065191  
AK065191.1 GI:32975209  
VERSION  
FLI CDNA; CAP trapper.  
KEYWORDS  
Oryza sativa (japonica cultivar-group)  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartioideae; Oryzeae; Oryza.



REFERENCE  
AUTHORS

1 The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team:, Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.

TITLE  
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
JOURNAL  
Science 301 (5631), 376-379 (2003)  
MEDLINE  
22752273  
PUBMED  
12869764

REFERENCE  
AUTHORS

2 (bases 1 to 2368)  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.

TITLE  
JOURNAL

Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT  
This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.  
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,

Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.

FEATURES  
source

1. .2368  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J013002E10"

ORIGIN

Alignment Scores:  
Pred. No.: 6.98e-163 Length: 2368  
Score: 2583.00 Matches: 508  
Percent Similarity: 83.74% Conservative: 69  
Best Local Similarity: 73.73% Mismatches: 106  
Query Match: 63.39% Indels: 7  
DB: 8 Gaps: 5

US-10-051-909-32 (1-800) x AK065191 (1-2368)

QY 109 IleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArg 128  
Db 2 CTTGGGCGACGATCATCACGCGTTCTCTGGAGCAGTGGCTGATTCTTTGGTAGGCGG 61

QY 129 ProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSer 148  
Db 62 CCCATGCTGATCGCGTCGGCTGTCTCTACTTTGTTAGTGGCTAGTGATGCTTTGGGCG 121

QY 149 ProAsnValTyrValLeuLeuAlaAargPheValAspGlyPheGlyIleGlyLeuAla 168  
Db 122 CCAATGTGTATGTGTGCTCTTGGCGAGGCTCATTTGACGGGTTCGGATCGGTTGGCT 181

QY 169 ValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeu 188  
Db 182 GTCACGCTTGTAACCATTTGTACATCTCTGAGACTGCCCGACGACATCAGAGGACTGCTA 241

QY 189 AsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPhe 208  
Db 242 AACACGTCGCCGAGTTCAGTGGGTCTGGAGGGATGTTCTCTTCATACTGCATGGTATTT 301

QY 209 GlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIlePro 228  
Db 302 GGCATGTCCCTCATGCCACAGCAGATTGGAGGATCATGTGTGGCGTTCTATCAATACCA 361

QY 229 SerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuVal 248  
Db 362 TCACTTATATACTTTTGCATTTGACCATCTTTTACTTTACCTGAATCGCCGAGGTGGCTCGTG 421

QY 249 SerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAsp 268  
Db 422 AGCAAAGGAAGAATGGCTGAGGCCAAGCGTGTGTGTGCAAGGCCTGCGTGAAGAGAAGAT 481

QY 269 ValSerGlyGluLeuSerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIle 288  
Db 482 GTTTCAGGAGAAATGGCCCTTCTCGTTGAAGGTCTGGGGSTTGGGAAAGACACAAAAATT 541

QY 289 GluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAsp 308  
Db 542 GAGGAATACATAATTGGACCTGATGATGAGCTTGCTGTATGAAGGGCTGGCTCCA---GAT 598

QY 309 LysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSer 328  
Db 599 CCAGAGAAGATCAAACTGTATGGTCTCTGAAGAAGGCTTATCGTGGGTGGCCCGTCTCTGT 658

QY 329 LysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetVal 348  
Db 659 CACGGGCAAAAGTGCACTTGGAAAGTGCATTAGGTCTCATCTCTCGTCATGGTAGTAGTGGTC 718

Qy 349 AsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsn 368  
Db 719 AGTCAGGGTAAGCCCTTGTGGATCCTTTTGTACCCCTTTTGGAAAGTGTCCATGAGAAG 778  
Qy 369 MetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPhe 388  
Db 779 ATGCCCTGAGATAATGGGAAGCATGCGGAGCACATTGTTTCCTAACTTTGGCAGCATGTTT 838  
Qy 389 SerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAsp 408  
Db 839 AGTGTGGCGGAACAGCAGCAAGCTAAAGGTGATTGGGATGCTGAGAGT---CAACGGGAG 895  
Qy 409 AspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerPro 428  
Db 896 GGTGAAGATTATGGATCAGACCATGGTGGGATGATGACATTGAAGATAC-CTCCAAAGCCCCA 954  
Qy 429 LeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle---ValHisHisGlyHis 447  
Db 955 CTTATTCTCGTCAAGCGACAAAGCTGGAAGGAAGGAGATCGCTGCACCTCATGGCAGT 1014  
Qy 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyVal 467  
Db 1015 ATAATGGTGCTGTGGGA-----AGAAGTAGTAGTCTCATGCAGGCGCGGAGGCAGTA 1068  
Qy 468 SerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGly 487  
Db 1069 AGCAGCATGGGCATTGGTGGGGATGGCAGTTGGCTTGGAATGGACTGAGAGAGAAGGT 1128  
Qy 488 GluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValPro 507  
Db 1129 GCAGATGGCGAAAGAAAGGTGGCTTCCAACGATATCTACTTGCATGAAGAGGGTGTGACA 1188  
Qy 508 GlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySer 527  
Db 1189 GGTGATCGCAGGGCTCTATACTGTCTATTCCT---GGAGGTGATGTCTCTCTGGTGGT 1245  
Qy 528 GluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAla 547  
Db 1246 GAGTTCGTCCAGGCAGCTGCTCTTGTGCAGCCAACTGCTCTTTACTCTAAGGAATTGATG 1305  
Qy 548 GluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySer 567  
Db 1306 GAGCAACGCCCTTGCTGGCCCTGCTATGGTGCATCCATCTCAGGCAGTGTCTAAAGGTCCA 1365  
Qy 568 ArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIle 587  
Db 1366 AAATGGGCAGACTTATTGAAACCTGGAGTGAAGCATGCTCTGTTTGTGGCATAGGGATA 1425  
Qy 588 GlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeu 607  
Db 1426 CAAATCCTGCAACAGTTTGTCTGGCATTAAATGGAGTTCTGTACTACATCCACAAATTCIT 1485  
Qy 608 GluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSer 627  
Db 1486 GAGCAAGCTGGTGTGTGTCTTCTTGCAAAACATTGGACTTAGCTCTCATCTGCATCT 1545  
Qy 628 IleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeu 647  
Db 1546 ATTCTTATTAGCGGACTGACAACCTTGCTGATGCTTCCCAGCATTTGGTATTGCTATGAGG 1605  
Qy 648 LeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAla 667  
Db 1606 CTCATGGATATGCTGGAAGAAGGTTTCTTCTCTTGCAACAATCCCTATCTATCTGATAGTA 1665  
Qy 668 SerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeu 687  
Db 1666 GCACTAGCTATCTTGATTCTGGTCAATATTCTGGATGTGGGGACCTGGTTTCATGCCTCA 1725  
Qy 688 LeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIle 707  
Db 1726 CTGTCCACAGTCAGTGTGCATACTCTACTTCTGCTTCTTTGTGTCATGGGGTTCGGGCCATT 1785  
Qy 708 ProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIle 727

Db 1786 CCAAACATTCTCTGTGCAGAGATTTTCCCGACCACCGTTCTGTGGCATCTGCATAGCCATC 1845  
Qy 728 CysAlaPheThrPheTrpIleGlyAspIleValThrTyrSerLeuProValMetLeu 747  
Db 1846 TGTGCCCTTAACATTCTGGATCGGTGATATCATTTGTACATACACCCCTCCCGTGTATGCTC 1905  
Qy 748 AsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPhe 767  
Db 1906 AACGCCATTGGACTCGCTGGAGTGTTTTGGAACTTACGCAGTGGTCTGCATACTGGCTTTC 1965  
Qy 768 ValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGlu 787  
Db 1966 CTGTTTGTCTTCATGAAGTGCCTGGAGACAAAGGCGCATGCCTCTTGAAGTCATCACCGAG 2025  
Qy 788 PhePheAlaValGlyAlaLysGlnAla 796  
Db 2026 TTCTTCTCTGTCGGAGCAAAGCAGGCC 2052  
RESULT 13  
AC073166  
LOCUS  
DEFINITION  
Oryza sativa chromosome 10 BAC OSJNBb0064P21 DNA linear PLN 28-AUG-2001  
complete sequence.  
AC073166  
AC073166.7 GI:12039314  
HTG.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
AUTHORS  
1 (bases 1 to 142114)  
Buell,C.R., Yuan,Q., Moffat,K.S., Hill,J.N., Burr,P.C., Hsiao,J.,  
Zismann,V., Pai,G., Bowman,C.L., Fujii,C.Y., VanAken,S.E.,  
Bowman,C.L., Craven,B., Utterback,T.R., Khalak,H., Feldblyum,T.V.,  
Quackenbush,J., White,O., Salzberg,S.L. and Fraser,C.M.  
Oryza sativa chromosome 10 BAC OSJNBb0064P21 genomic sequence  
Unpublished  
2 (bases 1 to 142114)  
Buell,R.  
Direct Submission  
Submitted (09-JUN-2000) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
3 (bases 1 to 142114)  
Buell,R.  
Direct Submission  
Submitted (05-JAN-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org  
4 (bases 1 to 142114)  
Buell,R.  
Direct Submission  
Submitted (28-AUG-2001) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org  
On Jan 5, 2001 this sequence version replaced gi:11968438.  
Address all correspondence to:rice@tigr.org

BAC clone OSJNBb0064P21 is from Oryza sativa chromosome 10  
The orientation of the sequence is from SP6 to T7 end of the BAC  
clone.  
Genes were identified by a combination of several methods: Gene  
prediction programs including Genscan and Genscan+ (Chris Burge,  
http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky,  
http://genemark.biology.gatech.edu/GeneMark/), and GeneSplicer  
(Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org),  
searches of the complete sequence against a peptide database and  
the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml).  
Annotated genes are named to indicate the level of evidence for  
their annotation. Genes with similarity to other proteins are named  
after the database hits. Genes without significant peptide  
similarity but with EST similarity are named as unknown proteins.  
Genes without protein or EST similarity, that are predicted by more  
than two gene prediction programs over most of their length are

annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

FEATURES	
source	Location/Qualifiers
	1. .142114
	/organism="Oryza sativa (japonica cultivar-group)"
	/mol_type="genomic DNA"
	/cultivar="Nipponbare"
	/sub_species="japonica"
	/db_xref="taxon:39947"
	/chromosome="10"
	/map="RZ500"
	/clone="OSJNBb0064P21"
gene	complement (3008. .3658)
	/gene="OSJNBb0064P21.16"
	/note="predicted by genscan+"
mRNA	complement (<3008. .>3658)
	/gene="OSJNBb0064P21.16"
CDS	complement (3008. .3658)
	/gene="OSJNBb0064P21.16"
	/codon_start=1
	/product="hypothetical protein"
	/protein_id="AAG46103.1"
	/db_xref="GI:12039315"
	/translation="MVAGDSQARLLVLALLRLLDPAAAAAEEPELFRHSDYRATVP
	ARGISVDFWAPFESNLTRLHEDLRAPRTPDLVLSGLWHMLHVTDAARYGDALA
	SVVDAAKSLRSLPVPPHFWLGLPLLVNHLNNTDAKKVHMNDTILQAYDLEVEQRG
	LLQRDGGPFLLLDVGKLSRGCGQQCTADGMHYDGDVYDAVLHMLNALVIESQRI"
repeat_region	complement (3574. .3596)
	/rpt_family="GC_rich"
repeat_region	complement (4269. .4297)
	/rpt_family=" (GGGA)n"
repeat_region	5613. .5633
	/rpt_family=" (GGAA)n"
gene	complement (5634. .6005)
	/gene="OSJNBb0064P21.1"
	/note="identical to histone H4 (H4C13) GI:168499 (Zea mays), (Plant Physiol Biochem 25, 235 247 (1987));EST D23040 from this gene"
mRNA	complement (<5634. .>6005)
	/gene="OSJNBb0064P21.1"
CDS	complement (5634. .5945)
	/gene="OSJNBb0064P21.1"
	/codon_start=1
	/product="histone H4"
	/protein_id="AAG46106.1"
	/db_xref="GI:12039318"
	/translation="MSGRGKGKGLGKGAKRHKVLRDNIQGITKPAIRRLARRGV
	KRISGLIYEETRVLKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQRTLYGFG
	G"
repeat_region	complement (7473. .7546)
repeat_region	7556. .7630
	/rpt_family=" (GGAGAA)n"
repeat_region	complement (8238. .8259)
	/rpt_family="AT_rich"
repeat_region	complement (8915. .8963)
	/rpt_family="AT_rich"
repeat_region	complement (11055. .11097)
	/rpt_family="AT_rich"
repeat_region	complement (11244. .11264)
	/rpt_family="AT_rich"
gene	14313. .18336
	/gene="OSJNBb0064P21.2"
	/note="identical to calcium dependent protein kinase
	GI:435466 (Oryza sativa) (Gene 129 (2), 183 189 (1993));
	EST C22538, C22540, AA750452, C99271, C22539 from this
	gene"
	join(14313. .15004,15338. .15481,15564. .15716,15894. .16009,
	16646. .16810,17112. .17333,17903. .18336)
mRNA	/gene="OSJNBb0064P21.2"

CDS	join(14323. .15004,15338. .15481,15564. .15716,15894. .16009,
	16646. .16810,17112. .17333,17903. .18025)
	/gene="OSJNBb0064P21.2"
	/codon_start=1
	/product="calcium-dependent protein kinase"
	/protein_id="AAG46110.1"
	/db_xref="GI:12039322"
	/translation="MGNSCQNGTYGNNYQNSNRFPQDRFASRYVDGNDTEDCYSGSSR
	ASLAGALRQGLNLKSPVLGYKTPNVRELYTLRELGGQFGKTYLCTEISTGCQYACK
	TILKSNLRCVSDIEDVRREIQIMHHLGQKNIVTIKDTYEDEQAVHIVMELCAGGELF
	SKIQRGHYSERKAAELIKIIVGIIETCHSHGVMHRDLKPENFLLLDADDEFSVKAID
	FGLSVFFRPGQVFREVVGSPYIAPEVLEKREYGEADITAGVILYVLLTGVPFPWAD
	TQSGIYEKVLDRIDFKSNRWPRISSAKDLIKMLCPYPSERLKAHEVLKHPWICDN
	GVATNRALDPSVLPRLKQFSAMNRLKLSLQIIAERLSEEEIVGLREMFKAMDTKNRS
	VVTFGELKGLKRYSSVFKOTEINDLMEAADTTSTINWEEFIAAAVSLNKIEREKHLM
	AAFTYFDKDGSGFITVDKLQKACMERNMEDTFLBEMILEVDQNNDGQIDYAEFVTMMQ
	SNFGLGWQTVESSLNVALREAPQVY"
repeat_region	complement (18675. .18732)
	/rpt_family="GC_rich"
repeat_region	complement (20605. .20634)
	/rpt_family="AT_rich"
gene	22393. .24891
	/gene="OSJNBb0064P21.6"
	/note="predicted by genemark hmm"
mRNA	<22393. .>24891
	/gene="OSJNBb0064P21.6"
CDS	22393. .24891
	/gene="OSJNBb0064P21.6"
	/codon_start=1
	/product="hypothetical protein"
	/protein_id="AAG46113.1"
	/db_xref="GI:12039325"
	/translation="MAGSGGEAMASSGLDLGEYLDLRPDAIHRRAAASVAIVRSGGDDGP
	RIVDGRDRRARSRRRLSSWRRPAGGETSVESRRYGFTAAAAAEVGGKTRAAAIW
	EWKPVRLSRIGKRRCGLLSVEVAGVRGVPASMDGLRLAVTVRKAETKDGAVQTMPA
	TVRGGCADFETLFRVRCNIYPAGGAGTGKPLKEPRRFVSVVPAEPGVRLGAHAVD
	VSSLVLSLQKSSEGRVRWFDTAVTLGKATGGELLLLKGFQLMDDAGLCCLYTOAAT
	EKVDVDDVSPALARAHNKNFSVARTSGPKLSADAAISPSMRAYKQLIDRLSVDEH
	GDPVTSILIPRLADDELSDVGLPEYEVVDKGVETKVEVHYHHRDVLKELDSIABEQ
	IEAIEALMTNGGKKSPPSPKKVDQKQCLDADEEMVTVEFLRKLEVVDDKGRCLKQPMT
	PRSEKKAAAAAPVVVDLPGGLGTAVQTRDGGFLVSMNPFDLPLEKGDAPPKLAMQ
	VSRPFVLPSSMAATGFDVLQMAAAGGADEVDRKVARLGGMDNLTGKTPEQVGFEGIA
	EAVIGRRRTEGASSAARSVLRVKLAAAVSDGRSERVATGIWTAADDPETLEEVIAT
	SLQKLEAMAVDALMIQAEMADEDAPFEVAPAGDATTVFDALVSPDEWSESRSDDR
	TVVAAIQLRDPSTRYEAVGAPMVAVVQSARLLGAAGSGGRFKVRSLHVGGVQLRCPV
	GGGAGGGRASWGAEROKLTAMQWALAHGPARAAGRRARTPTTTTTPSSQARQRPDVVW
	SLSSRVLAGMWLKTVRNPDVRVGATAAGSGGGGN"
repeat_region	complement (23766. .23788)
	/rpt_family="GC_rich"
repeat_region	complement (24616. .24639)
	/rpt_family="GC_rich"
repeat_region	complement (24693. .24714)
	/rpt_family="GC_rich"
repeat_region	24853. .24903
	/rpt_family=" (CGG)n"
repeat_region	complement (26149. .26194)
	/rpt_family="AT_rich"
repeat_region	complement (26861. .26905)
	/rpt_family="AT_rich"
repeat_region	complement (27502. .27537)
	/rpt_family="GC_rich"
repeat_region	complement (27628. .27663)
	/rpt_family="AT_rich"
repeat_region	complement (28100. .28132)
	/rpt_family="AT_rich"
repeat_region	complement (29892. .29929)
	/rpt_family=" (TA)n"
misc_feature	29893. .29928
	/note="with dinucleotide (AT) repeat (35 to 42 copies)"
	34449. .38345
gene	/gene="OSJNBb0064P21.3"
	/note="similar to sugar transporter GI:1495273
	(Arabidopsis thaliana); EST AA231788, D48171, C72446,



mRNA	AA749640 from this gene"					join(34449. .34603,34798. .34894,35286. .35852,35977. .36929,
Alignment Scores:						
Pred. No.:	1.06e-155	Length:	142114			
Score:	2514.00	Matches:	552			
Percent Similarity:	53.77%	Conservative:	75			
Best Local Similarity:	47.34%	Mismatches:	109			
Query Match:	61.69%	Indels:	431			
DB:	8	Gaps:	9			
US-10-051-909-32 (1-800) x AC073166 (1-142114)						
QY	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr	75			
Db	34816	ATGGCGGCGCGTGTGTCGCCATCGCGGCTCCATCGGCAACTTGTGTCAGGGCTGG	34875			
QY	76	AspAsnAlaThrIleAla-	81			
Db	34876	GATAATGCAACCATTCAGGTAAGTTCGTGGTTAGAGAAAAACGAATCTTTTTTTTCT	34935			
QY	81	-	81			
Db	34936	GCCATAGTTTGTGTGCCATGTGTCGTTATGTGACGCTAGGGCCCGTGTGTGGTGGCAATTAT	34995			
QY	81	-	81			
Db	34996	TGTGATACTTGTGGAGATCATGATGGATTTGGTGGAAATCATCTGCAGTAAAAATAGTTTT	35055			
QY	81	-	81			
Db	35056	AGTTGTTCTTCTCTGTTTCATGATATCACAAAAAAAACAACTGATATTTTTTTTCAGTA	35115			
QY	81	-	81			
Db	35116	GAATTTTGAGTCAATTATGAGTAGGATTTAACAAATGTACGGATCTCATGAATGATCTGT	35175			
QY	81	-	81			
Db	35176	ATTTATGGAACTTGATACTGTCTGCATCGTAAGGTTTCAGAAATTTGATTGATACCGATG	35235			
QY	82	-	AlaAlaValle	85		
Db	35236	GTTTGTGACTTCTTGTGACCCCTGTGGACTTGTGTGGGATGATCTTTTATCAGGTGCGGTACT	35295			
QY	85	uTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValse	105			
Db	35296	GTACATCAAGAAGGAATTCAACTTGCAGAGCGAGCCCTTATCGAAGGCCTGATCGTGGC	35355			
QY	105	rMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerI1	125			
Db	35356	CATGTCGCTCATTTGGGCGACGATCATCACGACGTTCTCTGGAGCAGTGGCTGATTTCTTT	35415			
QY	125	eGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMe	145			
Db	35416	TGGTAGGCGCCCATGCTGATCGCGTCGGCTGTCTCTACTTTGTAGTGGGCTAGTGAT	35475			
QY	145	tLeuTyrSerProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyI1	165			
Db	35476	GCTTTGGCGCCAAATGTGTATGTGTGCTCTTGGCAGGCTCATTGACGGGTTTCGGGAT	35535			
QY	165	eGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleAr	185			
Db	35536	CGGTTTGGCTGTACGCTTGTACCATTTGTACATCTCTGAGACTGCCCGACGACATCAG	35595			
QY	185	gGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCy	205			
Db	35596	AGGACTGCTAAACACGCTGCCGAGTTTCAGTGGGTCTGGAGGGATGTTCTTTTCATAC	35655			
QY	205	sMetValPheGlyMetSerLeuSerProSerProAspTyrArgIleMetLeuGlyValle	225			
Db	35656	CATGGTATTTGGCATGTCCCTCATGCCACAGCCAGATTGGAGGATCATGCTTGGCGTTCT	35715			

QY	225	uAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSerProAr	245
Db	35716	ATCAATACCATCACTATATATACTTTTGCAATTGACCATCTTTTACTTACCTGAATCGCCGAG	35775
QY	245	gTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgG1	265
Db	35776	GTGGCTCGTGAGCAAGGAAGAAATGGCTGAGGCCAAGCGTGTGTGCAAGGCCTGCGTGG	35835
QY	265	YlysAspAspValSer-----	270
Db	35836	AAGAGAAGATGTTTCAGGTTTGTATGTCTTACCCAGCATTTTTTCCCTGTGAGGATCTAGA	35895
QY	270	-----	270
Db	35896	AAGTGGAGAAACTATTTAATGCTATCAGTATAGGTATATTTATCTATGCGATGGAGTTTCT	35955
QY	271	-----GlyGluLeuSerLeuLeuGlyLeuGlyLeuGluValGlyG	284
Db	35956	CATGTTGAATTAACTGTGCAGGAGAAATGGCCCTTCTCGTTGAAGGTCTGGGGTTGGGA	36015
QY	284	lyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuV	304
Db	36016	AAGACACAAAAATTGAGGAATACATAATTGACCTGATGATGAGCTTGTGATGAAGGCG	36075
QY	304	alThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpI	324
Db	36076	TGGCTCCA---GATCCAGAGAAGATCAAACTGTATGGTCCCTGAAGAAGGCTTATCGI	36132
QY	324	leAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArGH	344
Db	36133	TTGCCCCGTCTCTTTCACGGGCAAAGTGCACTTGGAAAGTGCAATTAGGTCTCATCTCTC	36192
QY	344	isGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlyS	364
Db	36193	ATGTTAGTATGTCAGTCAGGTAAAGCCCTTGTGGATCCTGTGTCAACCTTTTGGAA	36252
QY	364	erValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnP	384
Db	36253	GTGTCCATGAGAAGATGCCCTGAGATAAATGGGAAGCATCGCGAGCACATTGTTTCTCTA	36312
QY	384	heGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluA	404
Db	36313	TTGGCAGCATGTTTAGTGTGGCGGAACAGCAGCAAGCTAAAGGTGATTGGGATGCTGAGA	36372
QY	404	snLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspA	424
Db	36373	GT---CAACGGGAGGTGAAGATTATGGATCAGACCATTGTTGGGATGACATTGAAGATA	36429
QY	424	snLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle--V	443
Db	36430	GCCTCCAAAGCCCACTTATTTCTCGTCAAGCGACAAAGCGTGGAAAGGAAGAGATCGCTG	36489
QY	443	alHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluG	463
Db	36490	CACCTCATGGCAGTATAATGGGTGCTGTGGGA-----AGAAGTAGTAGTCTCATGCAGG	36543
QY	463	lyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpS	483
Db	36544	GCGGGGAGGCAGTAAGCAGCATGGGCATTGGTGGGGGATGGCAGTTGGCTTGGAAATGGA	36603
QY	483	erGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisG	503
Db	36604	CTGAGAGAGAAGGTGCAGATGGCGAAAAAAGAGGTGGCTTCCAACGTATCTACTTGCATG	36663
QY	503	lnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspV	523
Db	36664	AAGAGGGTGTGACAGGTGATCGCAGGGGCTCTATACTGTCTATTGCCT---GGAGGTGATG	36720
QY	523	alPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheS	543
Db	36721	TTCTCTCTGGTGGTGGATTTCGTCAGGCAGCTGCTCTTGTGAGCCAAACCTGCTCTTACT	36780
QY	543	erLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValA	563

Db 36781 CTAAGGAATTGATGGAGCAACGCCCTTGCTGGCCCTGCTATGGTGATCCATCTCAGGCAG 36840  
Qy 563 laAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAlaLeuLeuV 583  
Db 36841 TTGCTAAAGGTCCTAAATGGGCAGACTTATTCGAACCTGGAGTGAAGCATGCTCTGTTTG 36900  
Qy 583 alGlyValGlyIleGlnIleLeuGln----- 591  
Db 36901 TTGGCATAGGGATACAAATCCTGCAACAGGTAAACCCCTTTTCTGAATTTATGTCAATGATC 36960  
Qy 591 ----- 591  
Db 36961 GAATGCATGTGCTATGGTTGAACCTAATTTCTAGTAGCTCCAGTATAAATTACTCTGAT 37020  
Qy 591 ----- 591  
Db 37021 TAATGTAACTGGTTAAAAATCAAAGCATCTTTCATGCAGTGTCCCTTCTAGGTTTCTAGCGGAG 37080  
Qy 591 ----- 591  
Db 37081 TCTCTCAATTAAAGCTCTTACTTACCTTACTCGAAAAACATGTTTAGCTGTTACGTGTTTT 37140  
Qy 591 ----- 591  
Db 37141 TCTTGCAATATATATTCATACAAGGACAATGAGCTAAACTGCAACAATAAATACACCTTTG 37200  
Qy 591 ----- 591  
Db 37201 ATGATTAGAATACAAATGCCATCAATGTTGTTTAGTTTTTACACTAGTAAAGTCAACCT 37260  
Qy 591 ----- 591  
Db 37261 TAACATCTTCAGTATTCATTATTAAAGAAAACCATTTGGGACTGGGCTATGTCACATGACC 37320  
Qy 591 ----- 591  
Db 37321 TGATTACGGGTATTCTATCTGAAGCACAGAAACAAATGTGCATGCTTTTATTTCCTG 37380  
Qy 591 ----- 591  
Db 37381 AATTTATTTGCTGAATATGGTTCTAGTTTCGCGAGAATAACTTTCAGAAACTTTTCACAGT 37440  
Qy 591 ----- 591  
Db 37441 GGCAATAATCATGTATATTCTCTTTTACTTCCCTCCACTCTTCCTATTTTCCCTTTGTAT 37500  
Qy 591 ----- 591  
Db 37501 ATTTTATATCAGTCTATTTCATGTTTTTACCCTCACTATAGCATACTGTAATGTTAATGTGCTAT 37560  
Qy 592 -----GlnPheAlaGlyIleAsnGlyValLeuTyThrProGln 605  
Db 37561 CAACCTTTCCATTATGTGCAGTTTGTGGCATTAATGGAGTTCTGTACTACACTCCACAA 37620  
Qy 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625  
Db 37621 ATTCTTGAGCAAGCTGGTGTGGTGTCTCTTGTGCAAAACATTGGAGCTTAGCTCCTCATCT 37680  
Qy 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645  
Db 37681 GCATCTATTCTTATAGCGGACTGACAACTTGCTGATGCTTCCAGCATTTGGTATTGCT 37740  
Qy 646 MetLeuLeuMetAspLeuSerGlyArgArg----- 655  
Db 37741 ATGAGGCTCATGGATATGCTGGAAGAGGTTAGACTTCATATTTTACATTTTTTTTGT 37800  
Qy 655 ----- 655  
Db 37801 GATTACATGCATCACTTCTAAGTGAAAGTATTATGCTAAACATTATTCTGCGTTGTTTCT 37860  
Qy 656 -----PheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672

Db 37861 CATTCAAGGTTTCTT-CTCCTTGCAACAATCCCTATCTCTGATAGTAGCACTAGCTATCTTG 37919  
Qy 673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692  
Db 37920 ATTCTGGTCAATATTCTGGATGTGGGACCATTGGTTTCATGCTCACTGTCCACAGTCAGT 37979  
Qy 693 ValIleValTyPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCy0 712  
Db 37980 GTCATACTCTACTTCTGCTTCTTTGTTCATGGGGTTCGGGCTTATTCCAAACATTCTCTGT 38039  
Qy 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732  
Db 38040 GCAGAGATTTTCCCGACCAACCGTTCGTGGCATCTGCATAGCCATCTGTGCCCTAACATTC 38099  
Qy 733 TrpIleGlyAspIleIleValThrTySerLeuProValMetLeuAsnAlaIleGlyLeu 752  
Db 38100 TGGATCGGTGATATCATTTGTGACATACACCTCCCCGTGATGCTCAACGCCATTGGACTC 38159  
Qy 753 AlaGlyValPheSerIleTyAlaValValCysLeuIleSerPheValPheValPheLeu 772  
Db 38160 GCTGGAGTGTTTTGAATCTACGCAGTGGTCTGCATACTGGCTTTCCTGTTCATG 38219  
Qy 773 LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGly 792  
Db 38220 AAGGTCCGGAGACAAAGGCGATGCCTCTTTGAAGTCATCACCGAGTTCTTCTCTGTCGA 38279  
Qy 793 AlaLysGlnAla 796  
Db 38280 GCAAGACAGGCC 38291

RESULT 14  
AE017116  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) DNA linear PLN 06-JUN-2003  
77 of the complete sequence.  
ACCESSION  
AE017116 AE016959  
VERSION  
AE017116.1 GI:31433303  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
1 (bases 1 to 300957)  
AUTHORS  
The Rice Chromosome 10 Sequencing Consortium  
CONSRMT  
In-depth view of structure, activity, and evolution of rice  
TITLE  
chromosome 10  
JOURNAL  
Science 300, 1566-1569 (2003)  
REFERENCE  
2 (bases 1 to 300957)  
AUTHORS  
Buell,C.R., Wing,R.A., McCombie,W.R., Messing,J. and Yuan,Q.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (05-MAY-2003) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA  
COMMENT  
This is the pseudomolecule for rice chromosome 10, which was  
constructed by resolving discrepancies between overlapping BACs,  
trimming the overlap regions, and linking the unique sequences to  
form a contiguous sequence. Genes in individual BAC clone were  
identified by a combination of several methods: Gene prediction  
programs, searches of the complete sequence against a peptide  
database and EST databases. Genes with similarity to other proteino  
are named after the database hits. Genes without significant  
peptide similarity but with EST similarity are named as unknown  
proteins. Genes without protein or EST similarity, that are  
predicted by more than two gene prediction programs over most of  
their length are annotated as hypothetical proteins. Genes  
encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy,  
http://genome.wustl.edu/eddy/tRNAscan-SE/).

FEATURES  
source  
1. 300957  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"

misc_feature	/db_xref="taxon:39947"	
	/chromosome="10"	
gene	<1. .12488	
	/note="Chromosome Sequence Derivation: nucleotide sequence in this region was derived from BAC clone OSJNBa0040D23 (GB:AC074196)."	
mRNA	<3310. .>4569	
	/locus_tag="OSJNBa0040D23.23"	
CDS	/note="similar to chloroplast nucleoid DNA binding protein CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"	
	<3310. .>4569	
gene	/locus_tag="OSJNBa0040D23.23"	
	3310. .4569	
mRNA	/locus_tag="OSJNBa0040D23.23"	
	3310. .4569	
CDS	/codon_start=1	
	/product="putative nucleoid DNA binding protein"	
gene	/protein_id="AAP54833.1"	
	/db_xref="GI:31433304"	
mRNA	/translation="MGRLAPMQLLVLCILSVTTCAAHGLRRGLDQQMRGRILADAT AAPGGAVVPLHWSGAHYVANFTIGTPPQAVSGIVDLSGELVMTQCAACRSSGCFKQE LPVFDPSANTYRAEQCGSPCKSKIPTNCSGDGECGYEAPSMFGDTFGIASSGDAIAI GNAEGRlafgcVcvvasdgsidgamdgsgfvglgRTPWSLVGQSNVTFsYCLALHGPg KsAlfLGASAKlAGAGKSNpPTPLlGQHASNTDDGSDPYTYVQLEGIKAGDVAVAA ASSGGGAITVlQLEtFRPLsYLPDAAYQALEKVVTAAALGSPsMANPPEFDLCFQNAa VSGVPDLVFTFOGATLTaQPSKYllGDGNGNGTVCLsILsTRLDsADDDGVSIlGSL LQENVHFLdLEKEtLSFEpADCSsLT"	
	<5409. .>5992	
mRNA	/locus_tag="OSJNBa0040D23.24"	
	3310. .5547,5811. .>5992	
CDS	/note="predicted by fgenesh"	
	join(<5409. .5547,5811. .>5992)	
gene	/locus_tag="OSJNBa0040D23.24"	
	3310. .5547,5811. .5992	
mRNA	/locus_tag="OSJNBa0040D23.24"	
	3310. .5547,5811. .5992	
CDS	/codon_start=1	
	/product="hypothetical protein"	
mRNA	/protein_id="AAP54834.1"	
	/db_xref="GI:31433305"	
gene	/translation="MTIGNLNSNIRICLTkTPCilKHPCtGAWMLRQRTSTKAAshCP TRGKRASKDSLI SAVLDNQNLNHGQTKAPLSTIIVARSGTFVARRNFPFKTRKKN MYAL"	
	<8062. .>9183	
mRNA	/locus_tag="OSJNBa0040D23.25"	
	3310. .9183	
CDS	/note="similar to chloroplast nucleoid DNA binding protein CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"	
	<8062. .>9183	
mRNA	/locus_tag="OSJNBa0040D23.25"	
	3310. .9183	
CDS	/codon_start=1	
	/product="putative nucleoid DNA binding protein"	
gene	/protein_id="AAP54835.1"	
	/db_xref="GI:31433306"	
mRNA	/translation="MRGRLLADAaVAGGGGVPIYWSQPLYMANLTIGTPPQPASAI IHLAGEFVWTQCSpcRRCFKQDlPLeNRSASsTYRPePCGTALCESVPASTCSGDGVC SYEVeTMFGDTSGIGGTDtFAIGTATASlAFGCAMDSNTKQLLGASGVVGLGRTPWSL VQOMNATAFsYCLAPHGAAGKKSALLLGASAKLAGKSAATtPLVNTSDSDSYMIHL EGIKFGDVIIAPPNGSVLVDTIFGVSFLVDAAFQAIKKAVTVAVGAAPMATPKPF DLCFPKAAAAGANSsSLPLPDVVLTfQGAAALtVPPSKMYDAGNGTVCLAMMSSAML NLtTELSIlGLRHQENIHFLFDLDKELsFEpADCSsLS"	
	<11356. .>12573	
mRNA	/locus_tag="OSJNBa0025B05.4"	
	3310. .12573	
CDS	/note="similar to chloroplast nucleoid DNA binding protein CND41 GB:BAA22813 GI:2541876 (Nicotiana tabacum)"	
	<11356. .>12573	
mRNA	/locus_tag="OSJNBa0025B05.4"	
	3310. .12573	
CDS	/locus_tag="OSJNBa0025B05.4"	
	3310. .12573	
gene	/codon_start=1	
	/product="putative chloroplast nucleoid DNA binding protein"	
mRNA	/protein_id="AAP54836.1"	
	/db_xref="GI:31433307"	
mRNA	/translation="MGRPvATLFLVlCFISVtACSLSEQATRGRLLAGVDATPPAAGGA	

misc_feature	VAVPIVLSsQGLYVANFTIGTPPQPVSaVVDLTgELVMTQCTPCQPCFEQDLPLFDPT KSSTFRGLPCGSHLCESIPeSSRNCTSDVCIYeAPTKAGDTGGKAGTDTFAIGAaKET LGFGCVMTDkRLKtIGPGSGIvGLGRTPWSLVtQMNVTAFsYCLAGKSSGAlFLGAT AKQLAGGKNsTPFVIKtSAGSSDNGSNPYVMVKLAGIKtGGAPLQAASSSGStVLLD TVSRASylADGAYKAlKAlTAAGVQpVAPSPPKPYDLCPFKAVAGDAPELVFTFDGG AAlTVPPANyLLASNGTVCLtIGSSASNLtGEGASlTGLSLQENVHVLFDLKEE TLsFKPADCSsLT"	
	12489. .23141	
gene	/note="Chromosome Sequence Derivation: nucleotide sequence in this region was derived from BAC clone OSJNBa0025B05 (GB:AC096782)."	
	complement(<16252. .>17307)	
mRNA	/locus_tag="OSJNBa0025B05.3"	
	3310. .4569	
CDS	/note="predicted by fgenesh"	
	complement(<16252. .>17307)	
mRNA	/locus_tag="OSJNBa0025B05.3"	
	3310. .4569	
CDS	complement(16252. .17307)	
	/locus_tag="OSJNBa0025B05.3"	
mRNA	/codon_start=1	
	/product="hypothetical protein"	
mRNA	/protein_id="AAP54837.1"	
	/db_xref="GI:31433308"	
mRNA	/translation="MVSAaAQaVVAaCVVlFVPMGLAGWHLsRNKVLFFSGAlFVSL AVGHlSPYlPSLPHLlLAAsFPHPHPISSSSSSSAaSSsCVPLlHRVSWADAGSGSV GRAWAPPSlASTCGLARlSDDASlLLNGSWVMAGDSQARlLVAlLLLDPAaA AAaEPElFRRHSDYRATVPARGISVDfVWAPfESNLTrLLHEDlRLAPRTPDVLVlGS GLWHMHVTDAARYGDAlASVVDAaKSLRSPlVPpPHMFwLGLPLlLVNHMLNTDAKK VHMNDlTlQAYDlEVEQGLlQDGGPFLLDVGKlSRGCGQCTADGmHYDGDVYDA VLHlMLNALVIESQQRi"	
	complement(<18878. .>19249)	
mRNA	/locus_tag="OSJNBb0064P21.1"	
	3310. .19249	
CDS	/note="similar to histone H4 (H4C13) GI:168499 (Zea mays), (Plant Physiol Biochem 25, 235 247 (1987));EST D23040 from this gene"	
	complement(<18878. .>19249)	
mRNA	/locus_tag="OSJNBb0064P21.1"	
	3310. .19249	
CDS	complement(18878. .19189)	
	/locus_tag="OSJNBb0064P21.1"	
mRNA	/codon_start=1	
	/product="histone H4"	
mRNA	/protein_id="AAP54838.1"	
	/db_xref="GI:31433309"	
mRNA	/translation="MSGRGKGKGLGKGAKRHKRVLRDNIQGITKPAIRRLARRGGV KRISGLIYEETRGVLKIFLENVIRDAVTYTEHARRKTVTAMDVVYALKRQRTLYGFG G"	
	<19395. .>20598	
mRNA	/locus_tag="OSJNBa0025B05.2"	
	3310. .20598	
CDS	/note="similar to B regulatory subunit of PP2A GB:AAH58900 GI:2160690 (Arabidopsis thaliana)"	
	join(<19395. .19473,19781. .19910,19992. .>20598)	
mRNA	/locus_tag="OSJNBa0025B05.2"	
	3310. .19473,19781. .19910,19992. .20598	
CDS	join(19395. .19473,19781. .19910,19992. .20598)	
	/locus_tag="OSJNBa0025B05.2"	
mRNA	/codon_start=1	
	/product="putative protein phosphatase 2A subunit"	
mRNA	/protein_id="AAP54839.1"	
	/db_xref="GI:31433310"	
mRNA	/translation="MTWLVTGRRDAIRGRVVLHRWMRMDGAGPPGDGGAGQDGRRL VPASGQCRRPPTRRRPRRPTCSTTrTLRNHVDRsFLAGlVALFAFEDSRERDLKYV YHQLYSKlTVERAFMQRIMAAAlLRFVYDTsPEEAERHCGVGELLEICGSINGFAVP LKEEHRAFMArVLLPLHRTFRWVtYHRQLsYCGSIrRSRHLPPPLAGDQMPEGAAHR RARGDRDPLsAPGLPAASAAAHrFVCFFIHfHNRfHFSQKKIDfYICTFN"	
	23142. .155358	
mRNA	/note="Chromosome Sequence Derivation: nucleotide sequence in this region was derived from BAC clone OSJNBb0064P21 (GB:AC073166)."	
	<27557. .>31580	
mRNA	/locus_tag="OSJNBb0064P21.2"	
	3310. .31580	
mRNA	/note="similar to calcium dependent protein kinase GI:435466 (Oryza sativa) (Gene 129 (2), 183 189 (1993)); EST C22538, C22540, AA750452, C99271, C22539 from this gene"	
	join(27557. .28248,28582. .28725,28808. .28960,29138. .29253,	



CDS  
29890.. .30054,30356.. .30577,31147.. .31580)  
/locus\_tag="OSJNBb0064P21.2"  
join(27567.. .28248,28582.. .28725,28808.. .28960,29138.. .29253,  
29890.. .30054,30356.. .30577,31147.. .31269)  
/locus\_tag="OSJNBb0064P21.2"  
/codon\_start=1  
/product="calcium-dependent protein kinase"  
/protein\_id="AAP54840.1"  
/db\_xref="GI:31433311"  
/translation="MGNSCQNGTYGNMYQNSNRFDNRFASRYVDGNDTDCYSGSSR  
ASLAGALRQGLNLKSPVLGYKTPNVRELYTLGRELGQGFQKTYLCTEISTGCCQYACK  
TILKSNLRCSVSDIEDVRREIQIMHLSGQKNIVTIKDTVEDEQAVHIVMELCAGGELF  
SKIQKRGHYSERKAAELIKIIVGIETCHSHGMHRDLKPENFLLLDADDEFVSXKAI

Alignment Scores:

Pred. No.: 3,13e-155 Length: 300957  
Score: 2514.00 Matches: 552  
Percent Similarity: 53.77% Conservative: 75  
Best Local Similarity: 47.34% Mismatches: 109  
Query Match: 61.69% Indels: 431  
DB: 8 Gaps: 9

US-10-051-909-32 (1-800) x AE017116 (1-300957)

QY	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75
Db	48060	ATGGCGGCGCGCTGCTGGTCGCCATCGCGCCTCCATCGGCAACTTGCTGCAGGGCTGG	48119
QY	76	AspAsnAlaThrIleAla-----	81
Db	48120	GATAATGCAACCATTCAGGTAAGTTCGTGGTTAGAGAAAAAACGAATCTTTT	48179
QY	81	-----	81
Db	48180	GCCATAGTTTGTGTGCCATGTCGTTATGTGACGCTAGGGCCCGTGT	48239
QY	81	-----	81
Db	48240	TGTGATACTTGTGGAGATCATGATGGATTTGGTGGAAATCATCTGCAGTAAAAATAGTTT	48299
QY	81	-----	81
Db	48300	AGTTGTTCTTCTGTTCATGATATCACAAAAAAACAAC	48359
QY	81	-----	81
Db	48360	GAATTTTGAGTCAATATGAGTAGGATTAAACAAATGTACGGATCTCATGAATGATCTGT	48419
QY	81	-----	81
Db	48420	ATTTATGGAACTTGATACTGTCTGCATCGTAAGGTTTCAGAAAATTGATTGATACCGATG	48479
QY	82	-----	85
Db	48480	GTTTGTGACTTCTGCTGACCCCTGTGGACTTGTGGGATGATCTTTATCAGGTGCGGTACT	48539
QY	85	uTyrIleLysLysGluPheGlnLeuGlnAenGluProThrValGluGlyLeuIleValse	105
Db	48540	GTACATCAAGAAAGAAATCAACTTGCAGAGCGAGCCCTTATCGAAGGCGCTGATCGTGGC	48599
QY	105	rMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerI1	125
Db	48600	CATGTCGCTCATTTGGGGCGACGATCATCACGACGTTCTCTGGAGCAGTGGCTGATTTCTTT	48659
QY	125	eGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMe	145
Db	48660	TGGTAGCGGGCCCATGCTGATCGCGTCGGCTGCTCTACTTTGTTAGTGGGCTAGTGAT	48719
QY	145	tLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyI1	165
Db	48720	GCTTTGGGCGCAAAATGTGTATGTGTTGCTTGTGGCAGGCTCATTCACGGGTTTCGGGAT	48779
QY	165	eGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleAr	185

Db	48780	CGGTTTGGCTGTACCGTTGTACCAATTGTACATCTCTGAGACTGCCCGACGGACATCAG	48839
QY	185	gGlyLeuLeuAasnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCy	205
Db	48840	AGGACTGCTAAACACCGTCCCGCAGTTCAGTGGGTCTGGAGGATGTTCCTTTTCATACTG	48899
QY	205	sMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValle	225
Db	48900	CATGGTATTGGCATGTCCCTCATGCCACAGCCAGATTGGAGGATCATGCTTGGCGTTCT	48959
QY	225	uAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProAr	245
Db	48960	ATCAATACCATCACTTATATACTTTTGCAATTGACCATCTTTTACTTACTGAATGCCCGAG	49019
QY	245	gTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValleuGlnLysLeuArgG1	265
Db	49020	GTGGCTCGTGAGCAAAGGAAGAAATGGCTGAGGCCAAGCGTGTGTGCAAGGCCCTCGGTGG	49079
QY	265	yLysAspAspValSer-----	270
Db	49080	AAGAGAAGATGTTTTCAGGTTTGTATGTCCTTATCCAGCATTTTTTCCCTGTGAGGATCTAGA	49139
QY	270	-----	270
Db	49140	AAGTGGAGAAAACATTTAATGCTATCAGTATAGGTATATATCTATGCGATGGAGTTTCT	49199
QY	271	-----GlyGluLeuSerLeuLeuGluGlyLeuGluValGlyG	284
Db	49200	CATGTTGTAATTAACGTGCAGGAGAAATGGCCCTTCTCGTTGAAGTCTGGGGGTGGGA	49259
QY	284	lyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuV	304
Db	49260	AAGACACAAAAATTGAGGAATACATAATTGGACCTGATGATGAGCTTGCTGATGAAGGGC	49319
QY	304	alThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpI	324
Db	49320	TGGCTCCA---GATCCAGAGAAGATCAAACTGTATGTCCTGAAGAAGGCTTATCGTGGG	49376
QY	324	leAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgH	344
Db	49377	TTGCCCCGTCCTGTTTACGGGCCAAAGTGACACTTGGAAAGTGCAATTAGGTCTCATCTCTCGTC	49436
QY	344	isGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlyS	364
Db	49437	ATGGTAGTATGGTCAGTCAGGGTAAGCCCTTGTGGATCCTGTGTGCACCCCTTTTGGAA	49496
QY	364	erValHisGluAasnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnP	384
Db	49497	GTGTCCATCAGAAAGATGCCTGAGATAATGGGAAGCATCGGAGACACATTTTTCCTAACT	49556
QY	384	heGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluA	404
Db	49557	TTGGCAGCATGTTTAGTGTGGCGGAACAGCAGCAGCAAGCTAAAGGTGATTTGGGATGCTGAGA	49616
QY	404	snLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspA	424
Db	49617	GT---CAACGGGAGGGTGAAGATTATGGATCAGACCATGCTGGGGATGACATTGAAGATA	49673
QY	424	snLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle---V	443
Db	49674	GCCTCCAAAAGCCCACTTATTTCTCGTCAAGCGACAAAGCGTGGAAAGGAAGGATCGCTG	49733
QY	443	alHiBHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluG	463
Db	49734	CACCTCATGGCAGTATAATGGGTGCTGTGGGA-----AGAAAGTAGTAGTCTCATGCAGG	49787
QY	463	lyGlyAspGlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrps	483
Db	49788	GCGGGAGGCAGTAAGCAGCATGGGCATTTGTTGGGGATGGCAGTTGGCTTGGAAATGGA	49847
QY	483	erGluLysGluGlyGluAasnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisG	503

Db 49848 CTGAGAGAGAAGGTGCAGATGGCGAAAAAGAGGTGGCTTCCAAACGATCTACTTGCATG 49907

Qy 503 InGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspV 523

Db 49908 AAGAGGGTGTGACAGGTGATCGCAGGGGCTCTATACTGTCTATGCTCT---GGAGGTGATG 49964

Qy 523 alPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheS 543

Db 49965 TTCTCTCTGGTGGTGGTTCGTCCAGGCAGCTGCTTGTTCAGCCAACTGCTCTTTTACT 50024

Qy 543 erLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValA 563

Db 50025 CTAAGGAATTGATGGAGCAACGCCTTGCTGGCCCTGCTATGGTGCAATCCATCTCAGGCAG 50084

Qy 563 laAlaLysGlySerArgTtpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuV 583

Db 50085 TTGCTAAAGGTCCAAAATGGGCAGACTTATTTCGAACCTGGAGTGAAGCATGCTCTGTTTG 50144

Qy 583 alGlyValGlyIleGlnIleLeuGln----- 591

Db 50145 TTGGCATAGGGATACAAATCCTGCAACAGGTAAACCTTTTCTGATTTTATGTATGTAC 50204

Qy 591 ----- 591

Db 50205 GAATGCATGTGCTATGGTTGAACCTTAATTTTCTAGTAGCTCCAGTATAAATTACTCTGAT 50264

Qy 591 ----- 591

Db 50265 TAATGTAACCTGGTTAAATCAAAAGCATCTTCATGCAGTGTCTCTTAGGTTTCTAGCGAG 50324

Qy 591 ----- 591

Db 50325 TCTCTCATTAAGCTCTTACTTACCTTACTCGAAAAACATGTTTAGCTGTGTACGTGTTTT 50384

Qy 591 ----- 591

Db 50385 TCTTGCATATATATTCATACAAAGGACAATGAGCTAAAACCTGCACAATAAATACACCTTTG 50444

Qy 591 ----- 591

Db 50445 ATGATTAGAAATACAAATGCCATCAATGTTTGTGTAGTTTTTACACTAGTAAAGTCAACCT 50504

Qy 591 ----- 591

Db 50505 TAACATCTTGCAGTATTCTATTATTAAGAAAACCATTGGGACTGGGCTATGTACATGACC 50564

Qy 591 ----- 591

Db 50565 TGATTACGGGTATTCTATCTGAAGCACAGAAACAATGTGCATGCAATCTTTTATTCTCTG 50624

Qy 591 ----- 591

Db 50625 AATTATTGCTGAATATGGTTCTAGTTTCGCAGAAATAACTTTCAGAAACTTTTCACAGT 50684

Qy 591 ----- 591

Db 50685 GGCATAATCATGTATATTCTTCTTTACTTCCCTCCACTCTTCTCCTATTCTCCCTTGAT 50744

Qy 591 ----- 591

Db 50745 ATTTTATATCAGTCTATTTCATGTTTACCTCACTATAGCATACTGAATGTTAATGTGCTAT 50804

Qy 592 -----GlnPheAlaGlyIleAenGlyValLeuTyrTyrThrProGln 605

Db 50805 CAACCTTTTCCATTATGTGCAGTTTGTCTGGCAATTAATGGAGTTCTGTACTACACTCCACAA 50864

Qy 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625

Db 50865 ATTCTTGAGCAAGCTGGTGTGGTGTCTTCTTCGCAACATTTAGCTCCTCATCT 50924

Qy 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645

Db 50925 GCATCTATTCTTATTAGCGGACTGACAACCTTGTCTGATGCTTCCCAGCATTTGGTATTGCT 50984

Qy 646 MetLeuLeuMetAspLeuSerGlyArgArg----- 655

Db 50985 ATGAGGCTCATGGATATGTCTGGAAAGAGGTAGACTTCATATTTTACATTTTTTTTGT 51044

Qy 655 ----- 655

Db 51045 GATTACATGCATCACTTCTAAGTGAAGATTAATATGCTAACCATTAATCTTTCGTTTCT 51104

Qy 656 -----PheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672

Db 51105 CATTTCAGGTTTCTT-CTCCTTGCAACAATCCCTATCCTGATAGTACTAGTATCTTG 51163

Qy 673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692

Db 51164 ATTCTGGTCAATATCTGGATGTGGGGACCATGGTTTCATGCCTCACTGTCCACAGTCAGT 51223

Qy 693 ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712

Db 51224 GTCATACTCTACTTCTGCTTCTTTGTGTCATGGGGTTCGGGGCTATTCCAAACATTCTCTGT 51283

Qy 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732

Db 51284 GCAGAGATTTTCCGACCACCGTTTCGTGGCATCTGCATAGCCATCTGTGCCCTAACATTTC 51343

Qy 733 TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752

Db 51344 TGGATCGGTGATATCATTTGTGACATACACCCCTCCCGTGTGATGTCACACGCCATTGGACT 51403

Qy 753 AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772

Db 51404 GCTGGAGTGTTCGAATCTACGCAGTGGTGTGCATACTGGCTTTCCTGTTTGTCTTCATG 51463

Qy 773 LysValProGluThrLysGlyMetProLeuGluValIleIleThrGluPhePheAlaValGly 792

Db 51464 AAGGTGCCGGAGACAAAGGCATGCCTCTTGAAGTCATCACCGAGTTCTTCTCTGTCCGA 51523

Qy 793 AlaLysGlnAla 796

Db 51524 GCAAAGCAGGCC 51535

RESULT 15

ATF23E12

LOCUS

DEFINITION Arabidopsis thaliana DNA chromosome 4, BAC clone F23E12 (ESSA project).

ACCESSION AL022604

VERSION AL022604.1 GI:3080406

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1

AUTHORS Bevan,M., Hilbert,H., Braun,M., Holzner,E., Brandt,A., Duesterhoeft,A., Hoheisel,J., Mewes,H.W., Mayer,K.F.X. and Schueller,C.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 86710)

AUTHORS EU Arabidopsis sequencing, project.

TITLE Direct Submission

JOURNAL

COMMENT Submitted (01-APR-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schuellemips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk

Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.

FEATURES

source

1. .86710

/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/variety="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="4"  
173. .1890  
/gene="F23E12.10"  
join(<173. .374,456. .919,1089. .1610,1678. .1890)  
/gene="F23E12.10"  
/codon\_start=1  
/product="hypothetical protein (fragment)"  
/protein\_id="CAA18726.1"  
/db\_xref="GI:3080407"  
/db\_xref="GOA:O65486"  
/db\_xref="UniProt/TrEMBL:O65486"  
/translation="FNAAVAGCFEAVESVLWPSSSTDSTSLPNTTSMVILSAVTASV  
VSEIGLSEPAFKVPDYDFRSPGELPLYLLGALCGLVSLALSRCTSSMTSAVDSLNK  
DAGIPKAVFPVMGGLSVGIITALVYPEVLYWGFQNVLDLLEKRPFKVGLSADLLQLVA  
VKIATAWCRASGLVGGYYAPSLFIGGAAGMAYGKFIGLALAQNPDFNLSILEVASPQ  
AYGLVGMAATLAGVCQPLTAVLLLFELTDYRIVPLPLGAVGMSSWITSGSKRQET  
RETKETRRKKSQEAQVQLTSSDSSSTNNLCEVESLCLDDSLNQSEELPKSIFVSEA  
MRTRPATVMSTSLAEALTRMLTEKQSCALIVDPDNIFLGILTLSDIQEFSKARKEGN  
NRPKDIFVNDICSRSGGCKVPWTVTPMDLLAAQTIMNKHLSHVAVVSGSIDAPRI  
HPVGVLDRRECITLRR"  
173. .374  
/gene="F23E12.10"  
/number=1  
375. .455  
/gene="F23E12.10"  
/number=1  
456. .919  
/gene="F23E12.10"  
/number=2  
920. .1088  
/gene="F23E12.10"  
/number=2  
1089. .1610  
/gene="F23E12.10"  
/number=3  
1611. .1677  
/gene="F23E12.10"  
/number=3  
1678. .1890  
/gene="F23E12.10"  
/number=4  
3124. .4428  
/gene="F23E12.20"  
join(3124. .3241,3323. .3465,3595. .3696,3945. .4107,  
4211. .4428)  
/gene="F23E12.20"  
/note="similarity to dihydroflavonol-4-reductase, Gerbera  
x sp., PIR2:S35189"  
/codon\_start=1  
/product="putative protein"  
/protein\_id="CAA18727.1"  
/db\_xref="GI:3080408"  
/db\_xref="UniProt/TrEMBL:O65487"  
/translation="MDQAKGKVCVTGASGFLASWLVKRLLEGYEVIGTVRDPGNEKK  
LAHLWKLEGAKERLRLVKADLMEEGSFDNAIMGCQGVFHTASPEILRPAIEGTLNVL  
SCRKNPSLKRVLTSSTSTVWYALSKTLAEQAANKFSEENGIDLTVLPSPFLVGPSP  
PDLCASTASDVGLLKGTEKFGWQMGYVHDDVARTHIVVFEHEAAQGRYICSSNV  
ISLEELVSFLSARYPSLPPIPKRFTDPLN"  
3124. .3241  
/gene="F23E12.20"  
/number=1  
3242. .3322  
/gene="F23E12.20"  
/number=1  
3323. .3465  
/gene="F23E12.20"  
/number=2  
3466. .3594  
/gene="F23E12.20"

/number=2  
3595. .3696  
/gene="F23E12.20"  
/number=3  
3697. .3944  
/gene="F23E12.20"  
/number=3  
3945. .4107  
/gene="F23E12.20"  
/number=4  
4108. .4210  
/gene="F23E12.20"  
/number=4  
4211. .4428  
/gene="F23E12.20"  
/number=5  
4939. .6163  
/gene="F23E12.30"  
complement(join(4939. .5013,5109. .5195,5307. .5342,  
5451. .5651,5725. .5811,6161. .6163))  
/gene="F23E12.30"  
complement(join(4939. .5013,5109. .5195,5307. .5342,  
5451. .5651,5725. .5811,6161. .6163))  
/gene="F23E12.30"  
/note="Contains Clathrin adaptor complexes small chain  
signature [VVYKRYASLYF]  
contains EST gb:T44219"  
/codon\_start=1  
/product="clathrin assembly protein AP19 homolog"  
/protein\_id="CAA18728.1"  
/db\_xref="GI:3080409"  
/db\_xref="GOA:O23685"  
/db\_xref="UniProt/TrEMBL:O23685"  
/translation="MIHFVLLVSRQGVRLTKWYSPYAKERSKVIRELSGVILNRGP  
KLCNPFVEWRGYKVYKRYASLYFCMCIDQEDNELEVLIIHHYVEILDYFGSVCELD  
LIFFNFHKAYYILDELLIAGELQESSKKTVARIISAQDLVEVAKEEASSISNIIAQAT  
NR"  
complement(4939. .5013)  
/gene="F23E12.30"  
/number=1  
complement(5014. .5108)  
/gene="F23E12.30"  
/number=1  
complement(5109. .5195)  
/gene="F23E12.30"  
/number=2  
complement(5196. .5306)  
/gene="F23E12.30"  
/number=2  
complement(5307. .5342)  
/gene="F23E12.30"  
/number=3  
complement(5343. .5450)  
/gene="F23E12.30"  
/number=3  
complement(5451. .5651)  
/gene="F23E12.30"  
/number=4  
complement(5652. .5724)  
/gene="F23E12.30"  
/number=4  
complement(5725. .5811)  
/gene="F23E12.30"  
/number=5  
complement(5812. .6160)  
/gene="F23E12.30"  
/number=5  
complement(6161. .6163)  
/gene="F23E12.30"  
/number=6  
7352. .7582  
/gene="F23E12.40"  
complement(7352. .7582)



CDS	/gene="F23E12.40"			
	complement (7352. .7582)			
	/gene="F23E12.40"			
	/codon_start=1			
	/product="hypothetical protein"			
	/protein_id="CAA18729.1"			
	/db_xref="GI:3080410"			
	/db_xref="UniProt/TREMBL:O65488"			
	/translation="MPLVRKNGVVMRIGSPMLHSENLVGDGVEAPPRGERDEIGEEV			
	GGETNVVSESCGEEGAQDAKTTPEGEADRI"			
exon	complement (7352. .7582)			
	/gene="F23E12.40"			
gene	/number=1			
	8233. .9045			
gene	/gene="F23E12.50"			
	complement (8233. .9045)			
CDS	/gene="F23E12.50"			
	complement (8233. .9045)			
	/note="similarity to DNA-binding protein PDI, Pisum sativum"			
	/codon_start=1			
	/product="putative protein"			
	/protein_id="CAA18730.1"			
Alignment Scores:				
Pred. No.:	3.27e-155	Length:	86710	
Score:	2502.00	Matches:	523	
Percent Similarity:	68.91%	Conservative:	91	
Best Local Similarity:	58.70%	Mismatches:	124	
Query Match:	61.40%	Indels:	153	
DB:	8	Gaps:	12	
US-10-051-909-32 (1-800) x ATF23E12 (1-86710)				
QY	52	LeuGluAspLys---MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsn	70	
Db	39649	ATAGAAGACGGAATTATGAGTGGAGCTGTGCTTGTGCTATTGCTGCTGTGTCGCAAC	39708	
QY	71	LeuLeuGlnGlyTrpAspAsnAlaThrIle-----	80	
Db	39709	TTGTTACAAGGATGGGATAACGCAACTATTGCAGGTGATTGATTATCTTCTGTATCCATA	39768	
QY	80	-----	80	
Db	39769	ATTCTGTTTATGAAATTTAGGATCGATAAAGTCAGGATTTTATGGATACAATTCACAGA	39828	
QY	81	-----	A1 81	
Db	39829	TCCTTTAGCTTATCTCTTTTGTCTGTGTGATGGATTTTAAAGTTTCTTTGGTACTCTGC	39888	
QY	81	aAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluG1	101	
Db	39889	AGGAGCTGTGTTGTACATAAAAAAGAGTTTAAATTTGGAGAGTAATCCATCAGTGGGAAGG	39948	
QY	101	yLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLe	121	
Db	39949	TCTAATTGGCGATGTCACCTTATTGGTGCTACTCTGATTACAACATGCTCTGGAGGGGT	40008	
QY	121	uSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSe	141	
Db	40009	AGCTGATTGGCTTGGTCGCCGTCCCATGCTAAATATTGCTCCTCAATTCTCTACTTTGTTGG	40068	
QY	141	rGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAs	161	
Db	40069	TTCTCTAGTAATGCATAGGTCTCCGAATGTTTATGTGTTGCTCTTAGGAAGTTTGTAGA	40128	
QY	161	pGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaPr	181	
Db	40129	TGGATTTGGGGTTGGTCTTGTTGGTCACACTGTTTCCTATTATATATCTGAGACTGCACC	40188	
QY	181	oSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMetPh	201	

Db	40189	ACCTGAGATTAGGGGACTGTTGAATACGCTACCGCAGTTTCACCTGGCTCTCGAGGGATGTT	40248	
QY	201	eLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMe	221	
Db	40249	CTTATCTTACTGTATGGTTTTCGGAATGTCTGTTGATGCCATCACCTAGCTGGAGATTGAT	40308	
QY	221	tLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuPr	241	
Db	40309	GCTTGGTGTCTTTTTCATCCCTTCCCTTGTCTTTTCTCCTCAGCGTCTTCTTCTTGCC	40368	
QY	241	oGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuG1	261	
Db	40369	CGAGTCCCCAAGGTGGCTCGTGAGCAAAGGTCGAATGCTTGAAGCAAAGCGGTTCTTCA	40428	
QY	261	nLysLeuArgGlyLysAspAspValSer-----	270	
Db	40429	GAGACTGCGTGGTCGCGAAGATGTCTGTTAGTCTTAAATCGTTCTGACTCTAATTCT	40488	
QY	271	-----GlyGluLeuS	274	
Db	40489	TTCAATGAAACCAATCTTTTCTATTGAATGTTCTTATGTGGCAATGTATAGGTGAGATGG	40548	
QY	274	erLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleG	294	
Db	40549	CTTTGTGGTTGAGGGTCTTGGAAATGGAGGTGAACAACCATAGAGGAATATATAAATTG	40608	
QY	294	lyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrL	314	
Db	40609	GTCCCGCGGATGAAGTTTACTGATGATCATGATATAGCTGTGGATAAGGATCAAATTAAAGT	40668	
QY	314	euTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetL	334	
Db	40669	TATATGGTGCAGAAAGAGGGCTGAGTTGGTTGCTAGGCCAGTCAAAGGA-----	40718	
QY	334	euGlySerValLeuSerLeuAlaSerArgHisGlySer--MetValAsnGlnSerValP	353	
Db	40719	--GGAAGCACATATGAGTGTGTTTGTCTCGCCATGGAAGTACAATGAGCAGGAGGCAAGGCT	40776	
QY	353	roLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaG	373	
Db	40777	CATTGATTGATCCTCTTGTGCACACTGTTTGGGAGCGTTCACGAGAAGATGCCCGGACACT-	40835	
QY	373	lyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspG	393	
Db	40836	--GGAAGCATGAGGAGTGCCCTTGTCCACATTTTGGGAGTATGTTTCAGTTGTGGAGGA	40893	
QY	393	lnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrA	413	
Db	40894	ATCAACCAAGACATGAAGATTGGGATGAAGAGAATCTTGTGGAGAAGGTGAGGATTATC	40953	
QY	413	laSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgG	433	
Db	40954	CATCCGAC---CATGGAGATGATTCTGAAGATGATCTTCAATCTCCGTTGATCTCACGTC	41010	
QY	433	lnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuS	453	
Db	41011	AAACGACAAGCATGGAG--AAAGACATGCCTCACACTGCTCATGGAACTCTTTCTACCT	41067	
QY	453	erMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleG	473	
Db	41068	TCAGACATGGAAGTCAAGTCAGGAGCTCAAGGGGAAGGACGGGTAGTATGGGGATTG	41127	
QY	473	lyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysG	493	
Db	41128	GAGGTGGATGGCAAGTGGCATGGAATGACGCGAAAGAGAGAAGATGAATCGGACAGAAAG	41187	
QY	493	luGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgLys	513	
Db	41188	AAGGTGGGTTTAAACGGATATATCTTGCATCAAGAAGGTTTCCCAGGATCTCGACGTGGCT	41247	
QY	513	erIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaA	533	
Db	41248	CAATTGTTTTCATTGCCTGGTGTGATGGAACCGGTGAG--GCAGATTTTGTACAAGCGCT	41304	

```
Qy      533  laAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerA 553
Db      41305 CTGCTTTGGTTAGCCACCAGCTCTTTATTTCCAAAGACCTTCTCAAAGAACAATCAATTG 41364

Qy      553  spAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuP 573
Db      41365 GTCCTGTATGGTACATCCATCCGAA--ACAACATAAAGGGTCAATTGGCATGATCTTC 41421

Qy      573  heGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln---- 591
Db      41422 ATGATCCTGGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAAAATACTTCAGCAGG 41481

Qy      591  -----
Db      41482 TAGTCTAGATAGCCACATCAAAATTAGTGAATCACAGAAGAATAATGCGGTTTATGCTA 41541

Qy      592  -----GlnPheAlaGlyIleAsnGlyValLeuT 601
Db      41542 ATTCTATGTTTTCCTGCTTCCTCTGGAAATGCAGTTCTCAGGCATCAACGGAGTTCCTT 41601

Qy      601  yrIyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyL 621
Db      41602 ACTACACACGCAAAATCCTTGAGCAGCGGGTGTCCGGATCCTACTATCGAACATGGGGA 41661

Qy      621  euSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProC 641
Db      41662 TTAGTTCCTCCTCAGCATCCTTACTTATTAAGTGCAATTGACAACCTTTGTGATGTTACCTG 41721

Qy      641  ysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg----- 654
Db      41722 CAATAGCTGTTGCAATGAGGCTCATGGATCTTTCTGGTCGAAGGTGATTAAACTTAAAT 41781

Qy      654  -----
Db      41782 TTCCTTTGTTCTTACAATACACATAAAACGAAACAAGCTAAAAATGCGTTATGAAATGCTT 41841

Qy      655  -----ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeu 669
Db      41842 TTATTTTATATTGTCAGGACCTTGCTTCTCACCACGATACCAATCCTGATAGCATCTCTA 41901

Qy      670  ValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSer 689
Db      41902 TTGGTTTTAGTAATCTCAAATCTTGTTACATGAACAGCATTTGTGCACGGGTCCTTATCA 41961

Qy      690  ThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsn 709
Db      41962 ACCGTAAGCGTGTGTCTCTACTTCTGCTTCTCGTAGGGGTTTCGGTCCTGCTCCAAAC 42021

Qy      710  ileLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAla 729
Db      42022 ATCCTCTGTTCAGAGATTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCATCTGCGCA 42081

Qy      730  PheThrPheTrpIleGlyAspIleValThrTyrSerLeuProValMetLeuAsnAla 749
Db      42082 CTCACCTTCTGGATCTGTGACATAATCGTCACTTACAGTCTCCCCCGTGTGCTCAAAATCC 42141

Qy      750  ileGlyLeuAlaGlyValPheSerIleTyrAlaValAlaValCysLeuIleSerPheValPhe 769
Db      42142 ATTGGACTAGCTGGTGTGTTTGGAAATGTAGCAATCGTATGTTGCATTTTCATGGGTCTTT 42201

Qy      770  ValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePhe 789
Db      42202 GTGTTTCATTAAAGTCCCGGAAACTAAAGGCATGCCACTTGAAGTCATCACAGAGTTCCTT 42261

Qy      790  AlaValGlyAlaLysGlnAlaAla 798
Db      42262 TCTGTTGGAGCTAGACAAGCTGAAGCT 42288
```

**This page blank (uspio)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 12:06:34 ; Search time 923.577 Seconds  
(without alignments)  
5127.663 Million cell updates/sec

Title: US-10-051-909-32  
Perfect score: 4075  
Sequence: 1 IRSGSLAVQTPTPDLRR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074037\_14040/app\_query.fasta\_1.1678  
-DB=N\_Geneseq\_16Dec04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909 @CGN\_1\_1\_891 @runat\_13042005\_074037\_14040 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :	N_Geneseq_16Dec04:*
1:	geneseqn1980s:*
2:	geneseqn1990s:*
3:	geneseqn2000s:*
4:	geneseqn2001as:*
5:	geneseqn2001bs:*
6:	geneseqn2002as:*
7:	geneseqn2002bs:*
8:	geneseqn2003as:*
9:	geneseqn2003bs:*
10:	geneseqn2003cs:*
11:	geneseqn2003ds:*
12:	geneseqn2004as:*
13:	geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB	ID Description
1	4075	100.0	2777	12 ADG47935 Adg47935 Corn Arab
2	4059	99.6	2908	12 ADM47932 Adm47932 Polynucle
3	3517	86.3	2824	6 ABK51962 Abk51962 Corn cont
4	3517	86.3	2824	8 ABX93198 Abx93198 cDNA enco
5	3517	86.3	2824	12 ADG47905 Adg47905 Corn Arab

6	2674	65.6	2601	6 ABK51965	Abk51965 Soybean c
7	2674	65.6	2601	8 ABX93201	Abx93201 cDNA enco
8	2674	65.6	2601	12 ADG47911	Adg47911 Corn Arab
9	2584	63.4	2190	6 ABZ13510	Abz13510 Arabidopo
10	2275	55.8	2205	6 ABZ14449	Abz14449 Arabidopo
11	2275	55.8	2205	6 ADG87978	Adg87978 A. thalia
12	1670.5	41.0	1692	6 ABK51966	Abk51966 Soybean c
13	1670.5	41.0	1692	8 ABX93202	Abx93202 cDNA enco
14	1670.5	41.0	1692	12 ADG47913	Adg47913 Soybean A
15	1424	34.9	1487	6 ABK51968	Abk51968 Wheat con
16	1424	34.9	1487	8 ABX93204	Abx93204 cDNA enco
17	1424	34.9	1487	12 ADG47917	Adg47917 Wheat Ara
18	1035	25.4	1412	12 ADJ39747	Adj39747 Plant cDN
19	943	23.1	1009	6 ABK51969	Abk51969 Wheat cDN
20	943	23.1	1009	8 ABX93205	Abx93205 cDNA enco
21	943	23.1	1009	12 ADG47919	Adg47919 Wheat Ara
22	620	15.2	751	12 ADJ41683	Adj41683 Plant cDN
23	616	15.1	1806	10 ADC07791	Adc07791 Rice DNA
24	600	14.7	778	12 ADJ41684	Adj41684 Plant cDN
25	587.5	14.4	1374	13 ADT48573	Adt48573 Bacterial
26	587.5	14.4	1518	6 ABL41880	Ab141880 Nucleotid
27	580	14.2	583	13 ACN56953	Acn56953 Cotton gy
28	562	13.8	870	6 ABK51964	Abk51964 Rice cDNA
29	562	13.8	870	8 ABX93200	Abx93200 cDNA enco
30	562	13.8	870	12 ADG47909	Adg47909 Rice Arab
31	552	13.5	659	12 ADJ41685	Adj41685 Plant cDN
32	548	13.4	644	12 ADJ42193	Adj42193 Plant cDN
33	539	13.2	1395	6 ABK73616	Abk73616 Bacillus
34	534	13.1	2127	12 ADM47936	Adm47936 Polynucle
35	533	13.1	1386	13 ADT48620	Adt48620 Bacterial
36	520.5	12.8	1826	3 AAC42332	Aac42332 Arabidopo
37	513	12.6	575	13 ACN46405	Acn46405 Cotton pr
38	505	12.4	1422	13 ADT48346	Adt48346 Bacterial
39	499	12.2	536	13 ACN56998	Acn56998 Cotton gy
40	496	12.2	547	13 ACN48128	Acn48128 Cotton pr
41	496	12.2	1853	6 ABK51973	Abk51973 Soybean c
42	496	12.2	1853	8 ABX93209	Abx93209 cDNA enco
43	496	12.2	1853	12 ADG47927	Adg47927 Soybean B
44	494.5	12.1	2017	6 ABK51972	Abk51972 Rice cont
45	494.5	12.1	2017	8 ABX93208	Abx93208 cDNA enco

ALIGNMENTS

RESULT 1	
ADG47935	
ID	ADG47935 standard; cDNA; 2777 BP.
XX	
AC	ADG47935;
XX	
DT	11-MAR-2004 (first entry)
XX	
DE	Corn Arabidopsis-like sugar transport protein cDNA #2.
XX	

Arabidopsis-like sugar transport protein;  
Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;  
corn; plant; gene; ss.

Zea mays.

Key Location/Qualifiers  
CDS 12..2414  
/\*tag= a  
/product= "Corn Arabidopsis-like sugar transport protein"  
/Note= "No start codon"  
/partial

US2002199217-A1.

26-DEC-2002.

17-JAN-2002; 2002US-00051909.

```
PR 24-APR-1998; 98US-0083044P.
PR 14-APR-1999; 99US-00291922.
PA (HELE/) HELENTJARIS T G.
XX Helentjaris TG;
XX WPI; 2004-040967/04.
DR P-PSDB; ADG47936.
XX
PT New isolated polynucleotide encoding a polypeptide having sugar transport
PT protein activity, for producing a transformed plant and for use as probes
PT in physical mapping.
XX
PS Claim 6; SEQ ID NO 31; 71pp; English.
XX
CC The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
CC transport proteins and their corresponding nucleic acid sequences. The
CC sequences of the invention are useful to transform a cell. These are also
CC useful to produce a transgenic plant. Probes derived from sequences
CC encoding sugar transport protein may be used for physical mapping. The
CC present sequence is corn Arabidopsis-like sugar transport protein cDNA.
XX
SQ Sequence 2777 BP; 667 A; 587 C; 701 G; 822 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2777
Score: 4075.00 Matches: 800
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-10-051-909-32 (1-800) x ADG47935 (1-2777)
QY 1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgArg 20
Db 12 ATTCGGAGCGGCTCTTGGCTTGAGTCCAAACGCCCTTCACCCCTGATCTGGACCGGAGG 71
QY 21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
Db 72 GAGCGGCTCTTCCGTCAGTTGTTCTTGTCTTGGCTGGGCTCTTCCGCCCTGCTTCGTGT 131
QY 41 SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
Db 132 TCTTCACAGGAGCGGCTGACCTCGACGATATCTTGGAGGACAAGATGTCGGGGCTGTT 191
QY 61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80
Db 192 CTTGTGCGCATAGTCGCCCTCCATCGGCAATCTATTGACGGGTGGGACAAATGCCACCATC 251
QY 81 AlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu 100
Db 252 GCAGCTGCTGTTCTGTATATAAAGAAGGAATTTCAATTGCAAAATGAGCCCACTGTGGAG 311
QY 101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyPro 120
Db 312 GGACTAATTGTGTCATGTCACTATCGCGCCACCATCGTTACTACATTCTCCGGGCCA 371
QY 121 LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140
Db 372 TTATCAGACTCGATTGGCCGACGCCCTATGCTTATCTCTCTTCAATTCTGTACTTCTTC 431
QY 141 SerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheVal 160
Db 432 AGCGGCCCTCATGCTATGGTCTCTTAATGTCTATGTCCTGCTGTTGGCACGTTTCGTA 491
QY 161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
Db 492 GATGGATTTGGTATTTGGCTTGCGTGTACGCTTGTGCTTGTACATTTTCAGAAATAGCC 551
QY 181 ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet 200
Db
```

```
Db 552 CCTTCGGAGATTAGAGGTTTGTGTAATACACTACCACAATTTCAGTGGATCAGGAGGAATG 611
QY 201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle 220
Db 612 TTCTTGTCTACTACTGTCATGGTGTGGGATGTCCTGTGCGCATCACCGATTGGAGAATT 671
QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240
Db 672 ATGCTTGGTGTCGCGATACCTTCATGTTCTCTTTGGTTTGACAATATTTTATCTT 731
QY 241 ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu 260
Db 732 CCTGAATCTCCAAGATGGCTCGTTAGCAAAAGTCCGATGGCAGAGGCAAAAAGGTGTTG 791
QY 261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGlyLeu 280
Db 792 CAAAAGTTACGGGGGAAAAGACGATGTCACAGGTGAATTGTCCCTTCTTCGAAGGGTTG 851
QY 281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300
Db 852 GAGGTTGGAGGAGACACTTCCATTGAAGAGTACATCATTTGGACCTGCCACCGAGGCAGCC 911
QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320
Db 912 GATGATCTTGTACTGACGGTGTATAAGGAACAAATCACACTTTATGGCCTGAAGAAGGC 971
QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340
Db 972 CAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATGCTTGGAAAGTGTCTTCTCTT 1031
QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360
Db 1032 GCATCTCGTCATGGGAGCATGGTGAACCAAGAGTGTACCCCTTATGGATCCGATTGTGACA 1091
QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380
Db 1092 CTTTTTGGTAGTGTCATGAGAAATATGCCCTCAAGCTGGAGGAAGTATGAGGAGCACATTG 1151
QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400
Db 1152 TTTCCAAACCTTGGAAAGTATGTTTCAGTGTACAGATCAGCATGCCAAAAATGAGCAGTGG 1211
QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp 420
Db 1212 GATGAAGAGAATCTTCATAGGGATGACGAGGAGTACGATCTGTATGGTCAGGAGGTGAC 1271
QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440
Db 1272 TATGAGGACAATCTCCATAGCCCATTTGCTGTCCAGGCAGGCAACAGGTGCGGAAGGGAAG 1331
QY 441 AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460
Db 1332 GACATTGTGCACCATGGTCCCGTACCGTGAAGTGTCTTGAGCATGAGAAAGGCAAAACCCCTCTTA 1391
QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp 480
Db 1392 GGGGAGGGTGGAGATGGTGTGAGCAGCAGTATCGTGGGGATGGCAGCTTGTCTTG 1451
QY 481 LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500
Db 1452 AAATGCTCAGAGAAGGAAGGTGAGAATGGTAGAAAGGAAGGTGGTTTCAAAAGAGTCTAC 1511
QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520
Db 1512 TTGCACCAAGAGGGAGTTCCTGGCTCAAGAAAGGGGCTCAATTGTTTTCACITCCCGGTGGT 1571
QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 1572 GGCGATGTTTGTAGGGGTACTGAGTTTGTACATGCTGCTGCTTTAGTAAGTCAGTCAGCA 1631
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560
Db 1632 CTTTTCTCAAAGGGTCTTGTGTAAACCAACGCAATGTCAGATGCTGCCATGGTTTCACCCATCT 1691
```





QY	101	GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyPro	120
DB	466	GGACTAATTGTGTCAATGTCACTTATCGGCGCCACCATCGTTACTACATTCTCGGGCCA	525
QY	121	LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe	140
DB	526	TTATCAGACTCGAATTGGCCGACGCCCTATGCTTATTCTCTTCAATTCTGTACTTCTTC	585
QY	141	SerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheVal	160
DB	586	AGCGGCTCATCATGCTATGGTCTCCTAATGTCTATGCTCTGCTGTGGCACGCTTCGTA	645
QY	161	AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla	180
DB	646	GATGGATTTGGTATTGGCTTGGCTGTCACGCTTGTGCCTTGTACATTTTCAGAAATAGCC	705
QY	181	ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMet	200
DB	706	CCTTCGGAGATTAGAGTTTGCTGGAATACACTACCACAAATTCAGTGGATCAGGAGGAATG	765
QY	201	PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle	220
DB	766	TTCTTGTCACTACTGCATGGTGTTTGGGATGTCCTCTGTCCTATCACCAGTTCGGAGAATT	825
QY	221	MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu	240
DB	826	ATGCTTGGTGTGCTCGCATACCTTCATTGTCTTCTTGGTTTGACAAATATTTTATCTT	885
QY	241	ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu	260
DB	886	CCTGAATCTCCAAGATGGCTCGTTAGCAAAAGTCGGATGGCAGAGGCCAAAAGGTGTG	945
QY	261	GlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeu	280
DB	946	CAAAAGTTACGGGGGAAGACGATGTCTCAGGTGAATTGTCCCTTCTTCTCGAAGGGTTG	1005
QY	281	GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla	300
DB	1006	GAGGTTGGAGGAGACACTTCCATTGAAAGAGTACATCATTTGGACCTGCCACCGAGGC	1065
QY	301	AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly	320
DB	1066	GATGATCTTGTTACTGACGGTGTATAAGGAACAAATCACACTTTATGGGCTTGAGAGGC	1125
QY	321	GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu	340
DB	1126	CAGTCATGGATTGCTCGACCTTCCAAGGGACCCAGCATGCTTGGAAGTGTGCTTTCTCTT	1185
QY	341	AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr	360
DB	1186	GCATCTCGTCATGGGAGCATGGTGAACCAGAGTGTACCCCTTATGGATCCGATTGTGACA	1245
QY	361	LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu	380
DB	1246	CTTTTGGTAGTGTCCATGAGAAATATGCCCTCAAGCTGGAGGAAGTATGAGGAGCACATTG	1305
QY	381	PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp	400
DB	1306	TTTCCAAACTTTGGAAGTATGTTTCAGTGTCCAGATCAGCATGCCAAATAATGAGCAGTGG	1365
QY	401	AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp	420
DB	1366	GATGAAGAGAAATCTTCATAGGGATGACGAGGAGTACGCATCTGATGGTGCAGGAGGTGAC	1425
QY	421	TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys	440
DB	1426	TATGAGGACAAATCTCCATAGCCCATTTGCTGTCCAGGCAGGCAACAGGTGCGGAAGGGAAG	1485
QY	441	AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu	460
DB	1486	GACATTGTGCACCATGGTCCACCGTGAAGTGTCTTTGAGCATGAGAAGGCCAAAGCCTCTTA	1545
QY	461	GlyGluGlyValAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp	480

Db	1546	GGGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGGTGGGGATGGCAGCTTGCTTGG	1605
QY	481	LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr	500
Db	1606	AAATGGTCAGAGAAAGGAAGGTGAGAAATGGTAGAAAGGAAGGTGGTTTCAAAGAGTCTAC	1665
QY	501	LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly	520
Db	1666	TTGCACCAAGAGGGAGTTCCCTGGCTCAAGAAAGGGCTCAATTGTTTCACTTCCCGGTGGT	1725
QY	521	GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla	540
Db	1726	GGCGATGTTCTTGAGGGTAGTGAGTTTGTACATGCTGCTGCTTTAGTAAGTCAGTCAGCA	1785
QY	541	LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaMetValHisProSer	560
Db	1786	CTTTTCTCAAAAGGTCTTGCTGAACACCGCATGTGAGATGCTGCCATGGTTCACCCATCT	1845
QY	561	GluValAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAla	580
Db	1846	GAGGTAGCTGCCAAAGGTTACGTTGGAAAGATTGTTTGAACCTGGAGTGAGCGGTGCC	1905
QY	581	LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu	600
Db	1906	CTGTTAGTCGGTGTGGAATTCTTGAGCAAGCTGGTGTGGCAGTTATTCTTTCCAAATTGGT	1965
QY	601	TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly	620
Db	1966	TACTATACCCCAAAATCTTGAGCAAGCTGGTGTGGCAGTTATTCTTTCCAAATTGGT	2025
QY	621	LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuPro	640
Db	2026	CTCAGCTCGGCATCAGCATCCATCTTGATCAGTCTCTCATACTACCTTACTAATGCTTCT	2085
QY	641	CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGly	660
Db	2086	TGCATTGGCTTTGCCATGCTGCTTATGGATCTTTCCGGAAGAAAGGTTTTTGTGCTAGGC	2145
QY	661	ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu	680
Db	2146	ACAAATCCAAATCTTGATAGCATCTCTAGTTATCTGGTGTGTGCCAATCTAATGATTG	2205
QY	681	GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe	700
Db	2206	GGTACACTAGCCCATGCTTTGCTCTCCACCCGTCAGTGTATCGTCTACTTCTGCTGCTTC	2265
QY	701	ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal	720
Db	2266	GTATATGGGATTTGGTCCCATCCCCAACATTTTATGTGCAGAGATCTTTCCAACAGGGTT	2325
QY	721	ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr	740
Db	2326	CGTGGCCCTCTGATTGCCATTTGTGCCCTTTACATTTCTGGATCGGAGATATCATCGTCACC	2385
QY	741	TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla	760
Db	2386	TACAGCCTTCCCTGTGATGCTGAATGCTATTGGACTGGCGGGTGTTCAGCATATATGCA	2445
QY	761	ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet	780
Db	2446	GTCGTATGCTTGATTTCCTTTGTGTTCTGCTCTTCCCTTAAGGTCCCTGAGACAAAGGGATG	2505
QY	781	ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla	800
Db	2506	CCCCCTTGAGGTTATTACCGAATCTTTTGCAGTTGGTGCGAAGCAAGCGGCTGCAAAAGCC	2565
RESULT 3			
ABK51962			
ID	ABK51962 standard; cDNA; 2824 BP.		
XX			
AC			
XX	ABK51962;		
XX			



```
Db      1498 GGGGGATGGCAGCTTGCTTGGAAATGGTCAGAGAAGGAGGTGAGATGGTAGAAAGGAA 1557
Qy      494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
Db      1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGAGTTCTTGGCTCAAGAAGGGGCTCA 1617
Qy      514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db      1618 ATTGTTTCACTTCCCGGTGGTGGCGATGTTCTTGAGGGTAGTGAGTTGTACATGCTGCT 1677
Qy      534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db      1678 GCTTTAGTAAGTCAGTCAGCACTTTTCTCAAAGGGTCTTGCTGAACCAACGCATGTCAGAT 1737
Qy      554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
Db      1738 GCTGCCATGGTTTACCCTCTGAGGTAGCTGCCAAAGGTTTCACGTTGGAAAGATTGTGTT 1797
Qy      574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
Db      1798 GAACCTGGAGTGAGGCGTGCCCTGTTAGTCGGTGTGGAATTCAGATCCTTCAACAGTTT 1857
Qy      594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db      1858 GCTGGAATAAACCAGGTGTTCTGTACTATATACCCACAAATCTTGAGCAAGCTGGTGGCA 1917
Qy      614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db      1918 GTTATTCTTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTC 1977
Qy      634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db      1978 ACTACCTTACTAATAGCTTCCTTGCAATGGCTTGGCATGCTGCTATGGATCTTTCCGGA 2037
Qy      654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db      2038 AGAAGGTTTTTGTGCTAGGCACAAATTCCAATCTTGATAGCATCTCTAGTTATCCTGGTT 2097
Qy      674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db      2098 GTGTCCAATCTAATTGATTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGTT 2157
Qy      694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db      2158 ATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGGTCCCATCCCCAACATTTTATGTGCA 2217
Qy      714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733
Db      2218 GAGATCTTTCCAAACCAGGGTTCGTGGCCTCTGTATTGCCATTTTGCTTTTACATTCTGG 2277
Qy      734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db      2278 ATCGGAGATATCATCGTCACCTACAGCCTTCCTGTGATGCTGAATGCTATTGGACTGGCG 2337
Qy      754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLys 773
Db      2338 GGTGTTTTTCAGCATATATGCAGTCGTATGCTTGAATTTCCCTTTTGTTGTTCTTCTTAAG 2397
Qy      774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db      2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTACCGAATTCCTTGCAAGTTGGTGCG 2457
Qy      794 LysGlnAlaAlaAlaLysAla 800
Db      2458 AAGCAAGCGGCTGCAAAAGCC 2478
RESULT 4
ABX93198
ID ABX93198 standard; cDNA; 2824 BP.
XX
AC ABX93198;
XX
DT 29-MAY-2003 (first entry)
```

```
XX      cDNA encoding corn sugar transport protein #1.
DE
XX      Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;
KW      plant sugar transport protein; carbohydrate transport; soybean;
KW      carbohydrate distribution; plant; gene; ss.
XX
OS      Zea mays.
XX
XX      US2002178468-A1.
PN
XX      28-NOV-2002.
PD
XX
XX      17-JAN-2002; 2002US-00051902.
PF
XX      24-APR-1998; 98US-0083044P.
PR      14-APR-1999; 99US-00291922.
XX
XX      (ALLE/) ALLEN S M.
PA      (HITZ/) HITZ W D.
PA      (KINN/) KINNEY A J.
PA      (TING/) TINGEY S V.
XX
XX      Allen SM, Hitz WD, Kinney AJ, Tingey SV;
PI
XX      WPI; 2003-340957/32.
DR
XX      P-PSDB; ABU08326.
DR
XX
PT      Novel plant sugar transport proteins and nucleic acid encoding the
PT      protein useful for producing transgenic plants having altered levels of
XX      sugar transport protein.
XX
PS      Claim 2; Page 13-15; 56pp; English.
SQ
CC      The present invention relates to the isolation of Arabidopsis thaliana-
CC      like or Beta vulgaris-like sugar transport proteins, and the
CC      polynucleotide sequences encoding them. The plant sugar transport
CC      proteins of the invention have been isolated from corn, rice, soybean,
CC      and wheat. The polypeptides of the invention may be used for altering the
CC      level of expression of a sugar transport protein in a host cell, by a
CC      transforming a host cell with a chimeric construct encoding all, or a
CC      portion of the sugar transport protein, in sense or antisense
CC      orientation. Particularly, the polypeptides may provide a means to
CC      control carbohydrate transport and distribution in plants. ABX93198-
CC      ABX93205 represent cDNA sequences encoding Arabidopsis thaliana-like
CC      sugar transport proteins
XX
SQ      Sequence 2824 BP; 644 A; 649 C; 745 G; 775 T; 0 U; 11 Other;

Alignment Scores:
Pred. No.:      4.45e-298      Length:      2824
Score:          3517.00      Matches:      684
Percent Similarity: 95.18%      Conservative: 27
Best Local Similarity: 91.57%      Mismatches: 34
Query Match:      86.31%      Indels:      2
DB:              8      Gaps:      2

US-10-051-909-32 (1-800) x ABX93198 (1-2824)
Qy      56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75
Db      238 ATGGGGGCGCGCTGATGGTCGCCCATCGCGCCTCTATCGGCAACTTGCTGCAGGGCTGG 297
Qy      76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db      298 GACAATGCGACAATTGCTGGAGCCGCTCTGTACATAAAGAGGAATTCACCTGCAGAGC 357
Qy      96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db      358 GAGCCTCTGATCGAAGGCCTCATCGTCGCCATGTTCTTCATTGGGGCAACAGTCATCACA 417
Qy      116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
```



Db 418 ACATCTCCGGGCCAAGGGCTGACTGCGTTGGTAGGAGGCCCATGCTGGTGCCTCGGCT 477

Qy 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 478 GTCCTCTACTTCGTCACTGGCTGGTGATGCTTTGGCGCCCAATTGTGTACATCTTGCTC 537

Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 538 CTCGCAAGGCTCATTTGATGGTTTCGGTATCGGTTTGGCGGTCAACATTGTTCTCTCTAC 597

Qy 176 IleSerGluIleAlaProSerGluIleArg---GlyLeuLeuAsnThrLeuProGlnPhe 194  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 598 ATCTCCGAAACTGCACCGCACAGANATTCTTGGGGCTGNNTNGAACAGTTGCCGAGTTC 657

Qy 195 SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 658 ATTGGGGTCAGNGGAGGATGTTCTCTCTCTACTGCATGGTGTGGGATGTCCTTCATG 717

Qy 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 718 CCCAAACCTGATTGGAGCTCATGCTTGGAGTTCTGTGATCCCGTCACTTATNACTTT 777

Qy 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 778 GGA CTGACTGCTTCTACTTGCCTGAATCACCAAGGTGGCTTGTNAGCAAGGATG 837

Qy 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 838 GCGGAGCGCAAGAGAGTGNLTGCAAAGGCTGCGGGAAGAGAGATGTCTCANGGGAGANG 897

Qy 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293  
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 898 GCTCTTCTAGTTGAAGGTTTGGGGTCGSTAAAGATACACGTAATTNAGAGTACATCA 957

Qy 294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 958 GGACCTGCCACCGAGGCCGATGATCTTGTAACTGACGGTGATAAGGAACAAATCACA 1017

Qy 314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1018 CTTTATGGGCTGAAGAGCCAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATG 1077

Qy 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1078 CTTGGAAGTGTGCTTTCTCTTGCACTCGTTCATGGGAGCATGGTGAACCCAGAGTGTACCC 1137

Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373

Db 1138 CTTATGGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGAAATATGCCTCAAGCTGGA 1197

Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1198 GGAAGTATGAGGAGCACATTGTTTCCAACTTTGGAAGTATGTTTCAGTGTACAGATCAG 1257

Qy 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1258 CATGCCAAAATGAGCAGTGGGATGAAGAGAAATCTTCATAGGGATGACGAGGATACGCA 1317

Qy 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1318 TCTGATGGTGAGGAGGTGACTATGAGGACAATCTCCATAGCCCAATTGCTGTCCAGGCAG 1377

Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1378 GCAACAGGTGCGGAAGGGAAGGACATTTGTGCACCATGGTTCACCGTGGAAAGTGCTTTGAGC 1437

Qy 454 MetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGly 473  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1438 ATGAGAAGGCAAGACCTCTTAGGGGAGGTGGAGATGGTGTGAGCAGCACTGATTCGGT 1497

Qy 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1498 GGGGGATGGCAGCTTGCTTGGAAATGGTCAGAGAAGGAAGGTGAGAAATGGTAGAAAGGAA 1557

Qy 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySer 513  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1558 GGTGGTTTCAAAAAGAGTCTACTTGCACCAAGAGGGAGTTCTTGGCTCAAGAAGGGGCTCA 1617

Qy 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1618 ATTGTTTCACTTCCCGTGGTGGCGATGTTCTTTGAGGGTAGTGAGTTGTACATGCTGCT 1677

Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1678 GCTTTAGTAAGTCAGTCAGCACCTTTTCTCAAAGGGTCTTGTCTGAACCAACGCATGTCAGAT 1737

Qy 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1738 GCTGCCATGGTTTCACCCATCTGAGGTAGCTGCCAAAGGTTTCAAGTTGGAAGATTTGTTT 1797

Qy 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1798 GAACCTGGAGTGAGGCGTGCCCTGTGTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTT 1857

Qy 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1858 GCTGGAATAAACGGTGTTCGTGACTATATACCCACAAATTCCTTGAGCAAGCTGGTGTGGCA 1917

Qy 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1918 GTTATTCTTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCCATCTTGATCAGTTCTCTC 1977

Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 1978 ACTACCTTACTAATGCTTCCTTCGATGGCTTGGCATGCTGCTATGGATCTTTCCGGA 2037

Qy 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 2038 AGAAGGTTTTTGTGCTGCTAGGCACAAATCCAAATCTTGATAGCATCTCTAGTTATCCTGGTT 2097

Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 2098 GTGTCCAATCTAAATTGATTTGGGTACACTAGCCCCCATGCTTTGCTCTCCACCATCAGTTT 2157

Qy 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 2158 ATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCA 2217

Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 2218 GAGATCTTTCCAAACCAGGGTTTCGTGGCCTCTGTATTGCCATTTGTGCCCTTTACATTTCTGG 2277

Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 2278 ATCGGAGATATCATCGTCACTACACGCTTCCTGTGATGCTGAATGCTATTGGACTGGCG 2337

Qy 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 2338 GGTGTTTTCAGCATATATGTCAGTCGTATGCTTGAATTTCCCTTTGTGTTCTTCCCTTAAG 2397

Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTACCGAATTCCTTTCAGTTGGTGGC 2457

Qy 794 LysGlnAlaAlaLysAla 800  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 2458 AAGCAACGGGCTGCAAAAGCC 2478

RESULT 5

ADG47905

ID ADG47905 standard; cDNA; 2824 BP.

XX

AC ADG47905;

XX

DT 11-MAR-2004 (first entry)

XX



Db 1318 TCTGATGGTGCAGGAGGTGACTATGAGGACAATCTCCATAGCCCAATTGCTGTCCAGGCAG 1377  
Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453  
Db 1378 GCAACAGGTGCGAAGGGAAGGACATTGTGCACCATGGTCAACCGTGAAGTGCTTTGAGC 1437  
Qy 454 MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 473  
Db 1438 ATGAGAAGGCCAAAGCCTCTTAGGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGGT 1497  
Qy 474 GlyGlyTrpGlnLeuAlaTirLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493  
Db 1498 GGGGATGGCAGCTTGCTTGGAATGGTCAGAGAAGGAAGGTGAGATGGTAGAAGGAA 1557  
Qy 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513  
Db 1558 GGTGGTTTCAAAAGAGTCTACTTGACCAAGAGGGAGTTCTTCGGTCAAGAAGGGGCTCA 1617  
Qy 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533  
Db 1618 ATTGTTTCACCTCCCGGTGGTGGCGATGTTCTTGAGGGTAGTGAGTTGTACATGCTGCT 1677  
Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553  
Db 1678 GCTTTAGTAAGTCAGTCAGCACTTTTCTCAAAGGGTCTTGCTGAACCAAGCATGTCAGAT 1737  
Qy 554 AlaAlaMetValHisProSerGluValAlaAlaAlaLysGlySerArgTrpLysAspLeuPhe 573  
Db 1738 GCTGCCATGGTTCACCCATCTGAGGTAGTGCCAAAGGTTACGTTGGAAAGATTTGTTT 1797  
Qy 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593  
Db 1798 GAACCTGAGTGAGGCGTGCCCTGTTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTT 1857  
Qy 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613  
Db 1858 GCTGGAAATAACGGTGTTCTGTACTATACCCCAAAATTCCTTGAGCAAGCTGGTGTGGCA 1917  
Qy 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633  
Db 1918 GTTATTCTTCCAAATTTGGTCTCAGCTCGGCATCGCATCCATCCTATCTGATCAGTTCTCTC 1977  
Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653  
Db 1978 ACTACCTTACTAATAGCTTCCTTGCAATGGCTTTGCCATGCTGCTTATGGATCTTCCGGA 2037  
Qy 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673  
Db 2038 AGAAGGTTTTTGCTGTAGGCACAATCCAAATCTTGATAGCATCTCTAGTTATCCTGGTT 2097  
Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693  
Db 2098 GTGTCCAATCTAATTGATTGGGTACACTAGCCCCATGCTTTGCTCTCCACCATCAGTGT 2157  
Qy 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713  
Db 2158 ATCGTCTACTTCTGCTGCTTCGTTATGGATTGGTCCCATCCCAACATTTTATGTGCA 2217  
Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTir 733  
Db 2218 GAGATCTTTCCAACACGAGGTTCGTGGCCTCTGTATTGCCATTTGTGCCTTACATTCTGG 2277  
Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753  
Db 2278 ATCGGAGATATCATCGTCACCTACAGCCTTCCTGTGATGCTGAATGCTATTGGACTGGCG 2337  
Qy 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773  
Db 2338 GGTGTTTTCAGCATATATGCAGTCGTATGCTTGATTTCCCTTTGTTGTTCTTCCTTAAG 2397  
Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793

Db 2398 GTCCTGTGAGACAAAGGGGATGCCCCCTTGAGGTTATTACCGAATTTCTTTCAGTTGGTGGC 2457  
Qy 794 LysGlnAlaAlaAlaLysAla 800  
Db 2458 AAGCAAGCGGCTGCAAAAGCC 2478  
RESULT 6  
ABK51965  
ID ABK51965 standard; cDNA; 2601 BP.  
XX  
AC ABK51965;  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE Soybean contig encoding A. thaliana-like sugar transport protein.  
XX  
KW Soybean; Arabidopsis thaliana-like sugar transport protein;  
KW carbohydrate transport; grain filling; annual field crop; plant; gene;  
KW ss.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 175..2388  
FT /\*tag= a  
FT /product= "Soybean Arabidopsis thaliana-like sugar  
FT transport protein"  
XX  
FN US6383776-B1.  
PD 07-MAY-2002.  
XX  
PF 14-APR-1999; 99US-00291922.  
XX  
PR 24-APR-1998; 98US-0083044P.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;  
XX  
DR WPI; 2002-453364/48.  
DR P-PSDB; AAU97204.  
XX  
PT New nucleic acid encoding plant sugar-transport proteins, useful for  
PT preparing transgenic plants with altered carbohydrate distribution.  
XX  
PS Claim 3; Col 35-38; 54pp; English.  
XX  
CC The present invention relates to the isolation of plant polynucleotide  
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein  
CC or Beta vulgaris-like sugar transport protein. The polynucleotide  
CC sequences are useful for altering the level of sugar transport proteins  
CC in plants, i.e. for control of carbohydrate transport and distribution in  
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,  
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and  
CC sugar transport. The polynucleotide sequences can also be used to isolate  
CC cDNA sequences and genes that encode homologues of the new proteins. The  
CC present sequence representing a contig assembled from various soybean  
CC cDNA clones encodes an Arabidopsis thaliana-like sugar transport protein  
XX  
SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.76e-224 Length: 2601  
Score: 2674.00 Matches: 522  
Percent Similarity: 82.20% Conservative: 92  
Best Local Similarity: 69.88% Mismatches: 115  
Query Match: 65.62% Indels: 18  
DB: 6 Gaps: 8

US-10-051-909-32 (1-800) x ABK51965 (1-2601)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTrp 75



Db 175 ATGAAAGGTGCCGTCCTTGTGTCTATTGCCGCTTCATTGGTAAATTCCTCCAAGGATGG 234  
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95  
Db 235 GATAATGCTACCATCGCCGGGCTAATGGTTACATTAAAGAAAGACCTTGCTTTGGGAACA 294  
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
Db 295 -----ACTATGGAAAGGCTTGTGTGGCATGTCCCTTGATTGGAGCAACGGTAATCACC 348  
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
Db 349 ACATGCTCTGGTCTCTATAGCGGATTGGCTCGGTGGCGACCCCATGATGATAATCTCATCT 408  
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155  
Db 409 GTGCTCTATTCTTGGGTGGTTTGGTGATGCTGTGGTCCCCAAATGTGTATGTGTGTGC 468  
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
Db 469 TTGGCGAGGCTACTTGATGGATTGGGATTGGCCTTGCTGTGACTCTTGTCCCCGGTCTAT 528  
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
Db 529 ATATCTGAACCGCGCGCTCTGAATAAAGGGGTCGTTGAATACGTTTCCTCAGTTTCAGT 588  
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
Db 589 GGCTCTGGAGGAATGTTTGTCTACTGTATGGTTTTGGCATGTTCATTGAGTCCCGCG 648  
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
Db 649 CCTAGCTGGAGGCTCATGCTTGGGTTCTGTCTATTCTCTCTGTATTTTGCATTG 708  
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
Db 709 ACCATTTTTCTTGCCCGAGTCTCCTCGGTGGCTGGTCAGCAAGAAAGGATGCTCGAG 768  
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275  
Db 769 GCTAAGAAAGGTGCTCCAAAGATTGCCGGAAGGGAGGATGTGTACGGCGAGATGGCATTG 828  
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
Db 829 CTGTTGAAGGTCTCGGATTGGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT 888  
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
Db 889 GCTGACGATGTGGTTCATGGTCAATGCAACAGAGAAAGATAAAATTCGATTATAT 948  
QY 316 GlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
Db 949 GGATCCCAAGCAGGCTTCTTGGTTATCAAAACCTGTCACTGGACAGAGTTCTATTGGC 1008  
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355  
Db 1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056  
QY 356 AspProIleValThrIlePheGlySerValHisGluAsnMetProGlnAlaGly----- 373  
Db 1057 GATCCTCTGGTGACACTGTTTGGTAGCATTCATGAGAAAGCTCCCGAGACAGGACAAGA 1116  
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
Db 1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGAAGCATGTTTCAGCACTGTCTGAGCCG 1176  
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413  
Db 1177 CATGCTAAAAATTGAACAATGGGATGAAGAAAGCTTACAAAGGAACGTCGAGGACTACATG 1236  
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433

Db 1237 TCAGATGCAACCCGTGGGGACTCCGATGATAAATTTGCACAGTCCTTTAATCTCACGCCAA 1296  
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452  
Db 1297 ACAACAAGCCTTGAA--AAAGACTTACCTCCTCCTCCTCCATGGCAGTATCCTTTGGC 1353  
QY 453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471  
Db 1354 AGCATGAGGCGTCACAGTAGTCTCTCATGCAAGGTCAGGTGAGCAAGGTGGTAGTACAGT 1413  
QY 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491  
Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAAATGGACTGATAAA--GGTGAGGATGGAAAA 1470  
QY 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511  
Db 1471 CAACAAGGAGGTTTTAAAGGATTTATTACATGAGGAGGAGTTTCTGCATCTCGTCGT 1530  
QY 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531  
Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTTGTCCAG 1572  
QY 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551  
Db 1573 GCTGCTGCCTTGGTAAGCCAAACCCGCTCTTTACTCCAAGAGCTTATTGATGGACACCCA 1632  
QY 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571  
Db 1633 GTTGGGCTGCAATGGTTCACCCCATCTGAGACAGCTTCAAAGGGCCCAAGTTGGAAAGCT 1692  
QY 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591  
Db 1693 CTTCTTGAACCAGGGGTTAAGCATGTCATTGGTTGTGGAGTTGGAATACAAATACTTTCAG 1752  
QY 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly 611  
Db 1753 CAGTTTTTCAGGATAAAATGGGGTTCTATATTACACACCTCAAATCCTTGAAGAGGCCGT 1812  
QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631  
Db 1813 GTTGAAGTTCTTCTTTCAGATATAGGCATTGGCTCAGAGTCGGCATCATTCCTTTATCAGT 1872  
QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651  
Db 1873 GCTTTCACAAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGGATGT 1932  
QY 652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671  
Db 1933 TCAGGCAGAAGGCAGTTGCTACTTACTACAAATCCCCGTCGTGATTGTGTCACTCATATT 1992  
QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691  
Db 1993 TTGGTCAATTGGAAGCCTGGTAAATTTTGGCAATGTCGCCCATGCAGCAATCTCAACAGTA 2052  
QY 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711  
Db 2053 TGC GTTGTGGTTTATTCTGTCTGCTTGTGTGATGGGTTATGGACCAATTCCAAACATCCTT 2112  
QY 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731  
Db 2113 TGCTCAGAGATTTTCCCACTAGGTCGTCGTCCTCTGCATTGCTATCTGTGCATTAGTG 2172  
QY 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751  
Db 2173 TTCTGGATTGGAGACATCATCATCACATCTCGTCGCTGTGATGCTCGGCTCTTTAGGA 2232  
QY 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771  
Db 2233 CTTGGTGGTGTATTGCGCCATTACGAGATTGTTTGTTCATCTCGTGGATATTGTGTTT 2292  
QY 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791  
Db 2293 TTGAAGGTTCCAGAAAAAAAGGGCATGCCCTTGAAGTCATCTCTGAATTCCTTTCTGTT 2352

QY 792 GlyAlaLysGlnAlaAlaAla 798  
|||||  
Db 2353 GGAGCAAAGCAGGCTGCTTCT 2373

RESULT 7  
ABX93201  
ID ABX93201 standard; cDNA; 2601 BP.  
XX  
AC ABX93201;  
XX 29-MAY-2003 (first entry)  
DT  
XX cDNA encoding soybean sugar transport protein #1.  
DE  
XX Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;  
KW plant sugar transport protein; carbohydrate transport; soybean;  
KW carbohydrate distribution; plant; gene; ss.  
XX  
OS Glycine max.  
XX  
FN US2002178468-A1.  
XX  
PD 28-NOV-2002.  
XX  
PF 17-JAN-2002; 2002US-00051902.  
XX  
PR 24-APR-1998; 98US-0083044P.  
PR 14-APR-1999; 99US-00291922.  
XX  
PA (ALLE/) ALLEN S M.  
PA (HITZ/) HITZ W D.  
PA (KINN/) KINNEY A J.  
PA (TING/) TINGEY S V.  
XX  
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;  
XX  
DR WPI; 2003-340957/32.  
DR P-PSDB; ABU08329.  
XX  
PT Novel plant sugar transport proteins and nucleic acid encoding the  
PT protein useful for producing transgenic plants having altered levels of  
PT sugar transport protein.  
XX  
PS Claim 2; Page 19-20; 56pp; English.  
XX

CC The present invention relates to the isolation of Arabidopsis thaliana-  
CC like or Beta vulgaris-like sugar transport proteins, and the  
CC polynucleotide sequences encoding them. The plant sugar transport  
CC proteins of the invention have been isolated from corn, rice, soybean,  
CC and wheat. The polypeptides of the invention may be used for altering the  
CC level of expression of a sugar transport protein in a host cell, by  
CC transforming a host cell with a chimeric construct encoding all, or a  
CC portion of the sugar transport protein, in sense or antisense  
CC orientation. Particularly, the polypeptides may provide a means to  
CC control carbohydrate transport and distribution in plants. ABX93198-  
CC ABX93205 represent cDNA sequences encoding Arabidopsis thaliana-like  
CC sugar transport proteins  
XX  
SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.76e-224 Length: 2601  
Score: 2674.00 Matches: 522  
Percent Similarity: 82.20% Conservative: 92  
Best Local Similarity: 69.88% Mismatches: 115  
Query Match: 65.62% Indels: 18  
DB: 8 Gaps: 8

US-10-051-909-32 (1-800) x ABX93201 (1-2601)

QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75  
||| |||||

Db 175 ATGAAAGGTGCGTCTTGTGTGCTATTGCGGCTTCCATTGGTAAATTTCTCCAAGGATGG 234  
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAen 95  
|||||  
Db 235 GATAATGCTACCATCGCCGGGCTAATGGTTACATTAAAGAAAGACCTTGCTTTGGGAACA 294  
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
|||||  
Db 295 -----ACTATGGAAAGGCTTGTGGTGGGCATGTCCCTGATTGGAGCAACGGTAATCACC 348  
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
|||||  
Db 349 ACATGCTCTGGTCTCTATAGCGGATTGGCTCGGTCCGCGACCCCATGATGATAATCTCATCT 408  
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155  
: : : : :  
Db 409 GTGCTCTATTCTTGGTGGTTTGGTGATGCTGTGGTCCCCAAATGTGTATGTGTGTGC 468  
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
|||||  
Db 469 TTGGCGAGGCTACTTGATGGATTGGGATTGGCCTTGCTGTGACTCTTGTCCCCGGTCTAT 528  
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
|||||  
Db 529 ATATCTGAACGGCGCGCTCTGAATAAGGGGGTCGTTGAATACGCTTCCTCAGTTCAGT 588  
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
|||||  
Db 589 GGCTCTGGAGGAATGTTTTTGTCTGCTACTGTATGGTTTTTGGCATGTATGAGTCCCGCG 648  
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
|||||  
Db 649 CCTAGCTGGAGGCTCATGCTTGGGGTTCTGTCTATTCTCTCTCTGTATTTTGCATTG 708  
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
|||||  
Db 709 ACCATTTTTTCTTGCCCGAGTCTCTCGGTGGTGGTCAGCAAAGGAAGGATGCTCGAG 768  
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
|||||  
Db 769 GCTAAGAAGGTGCTCCAAAGATTGCCGGAAGGGAGGATGTGTAGCGAGATGGCATTG 828  
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
|||||  
Db 829 CTGGTTGAAGTCTCGGATTGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT 888  
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
|||||  
Db 889 GCTGACGATGGGCTGATGGTCAATGAACATGCAACAGAGAAGATAAAAATTCGATTATAT 948  
QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
|||||  
Db 949 GGATCCCAAGCAGGCCCTTCTTGGTTATCAAAACCTGTCACTGGACAGAGTCTATTGGC 1008  
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355  
|||||  
Db 1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAAGCATGCCCTCATG 1056  
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373  
|||||  
Db 1057 GATCCTCTGGTGACACTGTTTGGTAGCATTCATGAGAAAGTCTCCCGAGACAGGACAAGA 1116  
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
|||||  
Db 1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGAAGCATGTTTCAGCACTGTGAGCCG 1176  
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413  
|||||  
Db 1177 CATGCTAAAAATTGAACAATGGGATGAAGAAAGCTTACAAAGGGAACGTGAGGACTACATG 1236  
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProIleLeuSerArgGln 433  
|||||  
Db 1237 TCAGATGCAAACCCGTGGGACTCCGATGATAATTTGCACAGTCTCTTAATCTCACGCCAA 1296

QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452  
Db 1297 ACAACAAGCCTTGAA---AAAGACTTACCTCCTCCTCCTCCATGGCAGTATCCTTGGC 1353  
QY 453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471  
Db 1354 AGCATGAGGCGTCACAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413  
QY 472 IleGlyGlyGlyTrpGlnLeuAlaTyrLysTrpSerGluLysGluGlyGluAsnGlyArg 491  
Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAAATGGACTGATAAA---GGTGAGGATGGAAAA 1470  
QY 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511  
Db 1471 CAACAAGAGGGGTTTAAAGGATTTATTATCATGAGGAGGAGTTTCTGCATCTCGTCGT 1530  
QY 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531  
Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTGTCCAG 1572  
QY 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551  
Db 1573 GCTGCTGCCTTGGTAAGCCCAACCCGCTCTTTACTCCAAGGAGCTTATTGATGGACACCCA 1632  
QY 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571  
Db 1633 GTTGGCCTGCAATGGTTCACCCATCTGAGACAGCTTCAAAGGGGCCAAGTTGGAAAGCT 1692  
QY 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591  
Db 1693 CTTCTTGAACACGAGGGTTAAGCATGCATTGGTTGTGGAGTTGGAAATACAAATACTTTCAG 1752  
QY 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly 611  
Db 1753 CAGTTTCAGGGATAAATGGGGTCTCTATATTACACACCTCAAATCCCTTGAAGAGGCCGT 1812  
QY 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631  
Db 1813 GTTGAAGTCTCTCTTTCAGATATAGGCATTGGCTCAGAGTCGGCATCATCTTATCAGT 1872  
QY 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651  
Db 1873 GCCTTCACAACCTTCTTGATGCTTCCCTGTATAGGCGTAGCCATGAAGCTCATGGATGTT 1932  
QY 652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671  
Db 1933 TCAGGCAGAAAGGCAGTTGCTACTTACTACAATCCCGCTGATTTGTCTCACTCATTATT 1992  
QY 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691  
Db 1993 TTGGTCATTGGAAGCCTGGTAAATTTTGGCAATGTGCGCCCATGCAGCAATCTCAACAGTA 2052  
QY 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711  
Db 2053 TGC GTTGGTTTATTCTCTGCTGCTTTGTGATGGGTATGGACCAATTCCAAACATCCTT 2112  
QY 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731  
Db 2113 TGCTCAGAGATTTTCCCACTAGGGTGGTGGCTCTGCAATTGCTATCTGTGCATTAGTG 2172  
QY 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751  
Db 2173 TTCTGGATTGGAGACATCATCATACATACATACTCGCTGCTGTGATGCTCGGCTCTTTAGGA 2232  
QY 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771  
Db 2233 CTTGGTGGTGTATTTCGCCATTTACGCAGTTGTTGTTTCATCTCGTGGATATTGTGTTT 2292  
QY 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791  
Db 2293 TTGAAGGTTCCAGAAACAAGGGCATGCCCTTGAAGTCACTCTCTGAATTCCTTTCTGTT 2352

QY 792 GlyAlaLysGlnAlaAlaAla 798  
Db 2353 GGAGCAAAAGCAGGCTGCTTCT 2373  
RESULT 8  
ADG47911  
ID ADG47911 standard; cDNA; 2601 BP.  
XX  
AC ADG47911;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Corn Arabidopsis-like sugar transport protein cDNA #2.  
XX  
KW Arabidopsis-like sugar transport protein;  
KW Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;  
KW corn; plant; gene; ss.  
XX  
OS Zea mays.  
XX  
FH  
FT Key Location/Qualifiers  
CDS 40..2388  
FT /\*tag= a  
FT /product= "Corn Arabidopsis-like sugar transport protein  
FT "  
XX  
XX US2002199217-A1.  
XX  
XX 26-DEC-2002.  
XX  
XX 17-JAN-2002; 2002US-00051909.  
XX  
XX 24-APR-1998; 98US-0083044P.  
XX 14-APR-1999; 99US-00291922.  
XX  
XX (HELE/) HELENTJARIS T G.  
XX  
XX Helentjaris TG;  
XX  
XX WPI; 2004-040967/04.  
DR P-PSDB; ADG47912.  
XX  
XX New isolated polynucleotide encoding a polypeptide having sugar transport  
PT protein activity, for producing a transformed plant and for use as probes  
PT in physical mapping.  
XX  
XX Disclosure; SEQ ID NO 7; 71pp; English.  
XX  
XX The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar  
CC transport proteins and their corresponding nucleic acid sequences. The  
CC sequences of the invention are useful to transform a cell. These are also  
CC useful to produce a transgenic plant. Probes derived from sequences  
CC encoding sugar transport protein may be used for physical mapping. The  
CC present sequence is corn Arabidopsis-like sugar transport protein cDNA.  
XX  
SQ Sequence 2601 BP; 648 A; 543 C; 644 G; 766 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 3.76e-224 Length: 2601  
Score: 2674.00 Matches: 522  
Percent Similarity: 82.20% Conservative: 92  
Best Local Similarity: 69.88% Mismatches: 115  
Query Match: 65.62% Indels: 18  
DB: 12 Gaps: 8  
US-10-051-909-32 (1-800) x ADG47911 (1-2601)  
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75  
Db 175 ATGAAAGGTGCCGCTCTTGTGCTATTGCGCTTCCATTGGTAATTTCTCCAAGGATGG 234  
QY 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95







Db 1240 GGATGGCAAGTGGCATGGAAATGGACGGAAGAGAGATGAATCGGACAGAAAGAA--- 1296

Qy 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIle 514

Db 1297 -----GAAGGTTTCCCAGGATCTCGACGTGGCTCAATT 1329

Qy 515 ValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 534

Db 1330 GTTTCATTGCCTGGTGGTGAACCGGTGAG--GCAGATTTTGTACAAGCGTCTGCT 1386

Qy 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554

Db 1387 TTGGTTAGCCCAACCACTCTTTATTCCAAAGACCTTCTCAAAGACATACAATTGGTCCT 1446

Qy 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574

Db 1447 GCTATGGTACATCCATCCGAA--ACAACATAAAGGGTCAATTGGCATGATCTTCATGAT 1503

Qy 575 ProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla 594

Db 1504 CCTGGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAAATACTTCAGCAGTTCTCA 1563

Qy 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614

Db 1564 GGCATCAACGGAGTTCTTTACTACACCGCAATCCTTGAGCAGCGGGTGTGCGGATC 1623

Qy 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634

Db 1624 CTACTATCGAACATGGGGATTAGTTCTCCTCAGCATCCTTACTTATAAGTGCAATTGACA 1683

Qy 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654

Db 1684 ACCTTTGTGATGTACCTGCAATAGCTGTGCAATGAGGCTCATGATCTTTCTGGTCGA 1743

Qy 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674

Db 1744 AGGACCTTGCTTCTCACCACGATACCAATCCTGATAGCATCTCTATTGGTTTAGTAATC 1803

Qy 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694

Db 1804 TCAAAATCTGTTACATGAACAGCATGTGACGCGGTCTTATCAACCGTAAGCGTTGTG 1863

Qy 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714

Db 1864 CTCTACTTCTGCTTCTTCGTGATGGGTTTCGGTCTCTGCTCCAAACATCCTCTGTCAGAG 1923

Qy 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734

Db 1924 ATTTTTCCAACTCGAGTCGCGGAATCTGCATCGCCATCTGCGCACTCACCTTCTGGATC 1983

Qy 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754

Db 1984 TGTGACATAATCGTCACTTACAGTCTCCCGTGTGCTGCTCAAAATCCATCGGACTAGCTGGT 2043

Qy 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774

Db 2044 GTGTTTGGAAATGTACGCAATCGTATGTTGCATTTCATGGGTCTTTGTGTTTCATTAAAGTC 2103

Qy 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794

Db 2104 CCGGAAACTAAAGGCATGCCACTTGAAGTCATCACAGAGTCTTTTCTGTTGGAGCTAGA 2163

Qy 795 GlnAlaAlaAla 798

Db 2164 CAAGCTGAAGCT 2175

RESULT 10

ABZ14449

ID ABZ14449 standard; DNA; 2205 BP.

XX

AC ABZ14449;

XX

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2254.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

KW Arabidopsis thaliana.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX WO200216655-A2.

PN 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

PF 24-AUG-2000; 2000US-0227866P.

XX 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.

PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

PI WPI; 2002-304127/34.

XX

DR Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.

PT Claim 144; SEQ ID NO 2254; 577pp + Sequence Listing; English.

XX

PS The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

XX

SQ Sequence 2205 BP; 521 A; 524 C; 521 G; 639 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.06e-189 Length: 2205

Score: 2275.00 Matches: 462

Percent Similarity: 75.89% Conservative: 111

Best Local Similarity: 61.19% Mismatches: 146

Query Match: 55.83% Indels: 36

DB: 6 Gaps: 14

US-10-051-909-32 (1-800) x ABZ14449 (1-2205)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGlnGlyTrp 75

Db 1 ATGAAGGAGCGACTCTCGTTGCTCTCGCCGCCACAATCGGCAATTCTTACAAGGATGG 60

Qy 76 AspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95

Db 61 GACAATGCCACCATTTGCTGGAGCTATGGTTTATATCAACAAGACTTGAATCTA----- 114

Qy 96 GluProThr---ValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114

Db 115 ---CCAACTCTGTTCAAGGTCCTGCTGCTATGCTATGCTATGCTCGGTCAACGGTCATC 171

Qy 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer 134

Db 172 ACGACTTGCTCAGGACCGATATCTGATTGGCTCGGCAGACGCCCATCATTTTATCA 231

Qy 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154



Db 232 TCAGTTATGTATTTTCGTCGTCCGGTTTGATAAATGTTGTGGTCTCCCAATGTCTATGTTCTG 291

Qy 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174

Db 292 TGCTTTGCTAGGCTTCTTAATGGGTTTGGTGCCGGGCTCGCGTTACACTGTCCCTGTT 351

Qy 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194

Db 352 TACATTCTGAACCGCTCCTCCGGAGATCAGAGGACAGTTAAATACTCTCCCTCAGTTT 411

Qy 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214

Db 412 CTTGGCTCTGGTGAATGTTTGTTCATACTCTATGCTGTTTTCACATATGTCCTGAGTGAC 471

Qy 215 SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234

Db 472 TCCCTAGCTGGAGAGCCATGCTCGGTGTCTCTCGATCCCTTCTCTTTATTGTTT 531

Qy 235 LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAla 254

Db 532 CTCACGGTGTTTTATTTGCCCGAGTCTCCTCGTTGGCTGGTTAGTAAAGGAATGGAC 591

Qy 255 GluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSer 274

Db 592 GAGGCTAAGCGAGTTCTTCAACAGTTATGTGCAGAGAAGATGTTACCGATGAGATGGCT 651

Qy 275 LeuLeuLeuGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGly 294

Db 652 TTACTAGTTGAAGGACTAGATATAGGAGGAGAAAAACAATGGAAGATCTCTTAGTAAC 711

Qy 295 ProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGlu---GlnIleThr 313

Db 712 TTGGAGGATCATGAAGGTGATGATACACTTGAACCCGTTGATGAGGATGGACAAATGCCG 771

Qy 314 LeuTyrGlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333

Db 772 CTTTATGGAACCCACGAGAAATCAATCGTACCTGTGTAGACCTGTCCAGAACAA --- 825

Qy 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353

Db 826 ---AATAGCTCACTTGGGCTACGCTCTCGCCACGGAAGCTTAGCAAAACCAAGCATGATC 882

Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373

Db 883 CTTAAAGATCCGCTCGTCAATCTTTTGGCAGTCTCCACGAGAAGATGCCAGAAGCAGGC 942

Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393

Db 943 GGAAACACTCGGAGTGGGATTTTCCCTCATTTCCGAAGCATGTTCAGTACTACTGCCGAT 1002

Qy 394 -----HisAlaLysAsnGluGlnTrpAsp-----GluGluAsnLeuHisArgAsp 408

Db 1003 GCGCCTCACGGTAAACCGGCTCATTTGGGAAAAGGACATAGAGAGCCATTACAACAAAGAC 1062

Qy 409 AspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp-----AsnLeu 425

Db 1063 AATGATGACTATGCGACTGATGATGGTGCGGGTGATGATGATGACTCGGACAAACGATTG 1122

Qy 426 HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis 445

Db 1123 CGTAGCCCTTAAATGTCGCGCCAGACCACAAGCATGGAC---AAGGATATGATCCACAT 1179

Qy 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAsp 465

Db 1180 CCTACAAGTGGAAGCACCTTTAAGCATGAGACGACACAGTACGCTTATGCAA---GGCAAC 1236

Qy 466 GlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLys 485

Db 1237 GCGGAAAGTAGCATGGGAATTGGTGGTGGTGGCATATGGGATATAGATACGAAAACGAT 1296

Qy 486 GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505

Db 1297 GAA-----TACAAGAGGTATATCTTAAAGAGATGGA 1329

Qy 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu 525

Db 1330 GCTGAA---TCTGCCGTGGCTCGATCATCTCTATTCCCGGAGGTCCGGAT-----GGT 1380

Qy 526 GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545

Db 1381 GGAGGCAGCTACATTACAGCTTCTGCCCTTGTAAGCAGATCTGTTCTTGGTCCTAAATCA 1440

Qy 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565

Db 1441 -----GTTCATGGATCCGCCATGGTTCCCCCGAGAAAAATTGCTGCCTCT 1485

Qy 566 GlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyVal 585

Db 1486 GGACCACTCTGGTCTGCTCTTCTTGAACCTGGTGTAAAGCGTGCCTTGGTTGGTGTC 1545

Qy 586 GlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605

Db 1546 GGCAATTCAAATACTGCAGCAGTTTTCAGGTATCAATGGAGTTCTCTACTACACTCCTCAG 1605

Qy 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625

Db 1606 ATTTCTGAACGGCTGGCGTAGATATTCTTTTCGAGCCTCGGACTAAGTTCCATCTCT 1665

Qy 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645

Db 1666 GCGTCAATCCTCATCAGCGGTTTAAACAACATTACTCATGTCCCAGCCATTGTCTGTGCC 1725

Qy 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeu 665

Db 1726 ATGAGACTCATGGATGTATCCGGAAGAAAGGTCAATTACTTCTCTGGACAATCCCAGTTCTC 1785

Qy 666 IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685

Db 1786 ATTGCTCTCACTTGTCTCCTTGTTCATCAGCGAGCTCATCCACATCAGCAAAAGTCGTGAAC 1845

Qy 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705

Db 1846 GCAGCACTCTCCACAGGTTGTGTGCTGTCTACTTCTGTCTTCTCGTATGGGTACGGT 1905

Qy 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725

Db 1906 CCCATTCCAAACATCCTCTGTCTTGAAATCTTCCCAACAAGAGTCCGTGGTCTCTGCATC 1965

Qy 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal 745

Db 1966 GCCATATGTGCTATGGTCTTTTGGATTGGAGACATTATTGTACAGTACTCACTTCCCGTT 2025

Qy 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765

Db 2026 CTCCTCAGCTCGATCGGACTAGTTGGTGTTTTTCAGCAATTACGTCGCGTTTTCGTTATC 2085

Qy 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785

Db 2086 TCATGGATCTTCGTTTACATGAAAGTCCCGGAGACTTAAAGGCATGCCCTTTGGAAGTTATC 2145

Qy 786 ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800

Db 2146 ACAGACTACTTTGCCTTTTGGAGCT---CAAGCTCAAGCTTCTGCT 2187

RESULT 11

ADG87978

ID ADG87978 standard; cDNA; 2205 BP.

XX

AC ADG87978;

XX

DT 22-APR-2004 (first entry)

XX

DE A. thaliana RPP4-upregulated pathogen infection-related gene #420.

XX

KW Pathogen infection-related gene; plant; Peronospora parasitica;

KW defence mechanism; RPP4; pathogen resistance; transgenic plant; oomycete;



Qy 446 GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAsp 465  
Db 1180 CCTACAAGTGAAGCACTTAAGCATGAGACGACACAGTACGCTTATGCAA--GGCAAC 1236  
Qy 466 GlyValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLys 485  
Db 1237 GCGCAAAGTAGCATGGGAATTGGTGGTTGGCATATGGATATAGATACGAAAACGAT 1296  
Qy 486 GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly 505  
Db 1297 GAA-----TACAAGAGGTATTATCTTAAAGAAGATGGA 1329  
Qy 506 ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu 525  
Db 1330 GCTGAA---TCTCGCCGTGGCTCGATCATCTATTCCCGGAGGTCGGAT-----GGT 1380  
Qy 526 GlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545  
Db 1381 GGAGGCAGCTACATTACGCTTCTGCCCTTGTAAAGCAGATCTGTTCTTGGTCTTAATCA 1440  
Qy 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565  
Db 1441 -----GTTCATGGATCCGCCATGGTTCCTCCCGAGAGAAATGCTGCCTCT 1485  
Qy 566 GlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyVal 585  
Db 1486 GGACCACTCTGGTCTGCTCTTTGAACCTGGTGTAAAGCGTGCCTTGGTGTGTC 1545  
Qy 586 GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln 605  
Db 1546 GGCATTCAAATACTGCAGCAGTTTTCAGGTATCAATGAGATTCTCTACTACACTCCTCAG 1605  
Qy 606 IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer 625  
Db 1606 ATTCTCGAACGGGCTGGCGTAGATATTCTTCTTCGAGCCTCGGACTAAGTCCATCTCT 1665  
Qy 626 AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla 645  
Db 1666 GCGTATTCTCTCATCAGCGGTTAAACAACATTACTCATGCTCCAGGCCATTGCTTGCC 1725  
Qy 646 MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeu 665  
Db 1726 ATGAGACTCATGGATGTATCCGGAAGAAGGTCAATTACTTCTTGGACAATCCCAGTTCTC 1785  
Qy 666 IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis 685  
Db 1786 ATTGTCTCACTTGTGCTCTCTCATCAGCGAGCTCATCCACATCAGCAAAGTCGTGAAC 1845  
Qy 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705  
Db 1846 GCAGCACTCTCCACAGGTTGTGTCGTCTCTACTTCTCTTTCGTGATGGGTTACGGT 1905  
Qy 706 ProfileProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725  
Db 1906 CCCATTCCAAACATCCTCTGTCTGAAATCTTCCCAACAAGAGTCCGIGTCTCTGCATC 1965  
Qy 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal 745  
Db 1966 GCCATATGTGCTATGGTCTTTTGGATTGGAGACATTATTGTACGTACTCACTTCCCGTT 2025  
Qy 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle 765  
Db 2026 CTCCTCAGCTCGATCGGACTAGTTGGTGTTTTTCAGCATTTTACGTCGCGGTTGCGTTATC 2085  
Qy 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785  
Db 2086 TCATGGATCTTCGTTTACATGAAAGTCCCGGAGACTAAAGGCATGCTCTTGGAGTTATC 2145  
Qy 786 ThrGluPhePheAlaValGlyValAlaLysGlnAlaAlaLysAla 800  
Db 2146 ACAGACTACTTTGCCTTTGGAGCT---CAAGCTCAAGCTTCTGCT 2187

RESULT 12  
ABK51966  
ID ABK51966 standard; cDNA; 1692 BP.  
XX AC ABK51966;  
XX 27-AUG-2002 (first entry)  
DE Soybean cDNA clone ssl.pk0022.f1 encoding sugar transport protein.  
XX KW Soybean; Arabidopsis thaliana-like sugar transport protein;  
KW carbohydrate transport; grain filling; annual field crop; plant;  
KW clone ssl.pk0022.f1; gene; ss.  
XX Glycine max.  
XX FH Key Location/Qualifiers  
CDS 9..1469  
FT /\*tag= a  
FT /partial  
FT /product= "Portion of a soybean Arabidopsis thaliana-like  
FT sugar transport protein"  
FT /note= "This sequence lacks a start codon"  
XX PN US6383776-B1.  
XX 07-MAY-2002.  
XX 14-APR-1999; 99US-00291922.  
XX 24-APR-1998; 98US-0083044P.  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;  
XX WPI; 2002-453364/48.  
XX P-PSDB; AAU97205.  
XX New nucleic acid encoding plant sugar-transport proteins, useful for  
XX preparing transgenic plants with altered carbohydrate distribution.  
XX Claim 3; Col 41-44; 54pp; English.  
XX The present invention relates to the isolation of plant polynucleotide  
XX sequences encoding an Arabidopsis thaliana-like sugar transport protein  
XX or Beta vulgaris-like sugar transport protein. The polynucleotide  
XX sequences are useful for altering the level of sugar transport proteins  
XX in plants, i.e. for control of carbohydrate transport and distribution in  
XX plant cells, e.g. during grain filling of annual field crops (e.g. corn,  
XX rice, soybeans, and wheat), and, for studying carbohydrate flows and  
XX sugar transport. The polynucleotide sequences can also be used to isolate  
XX cDNA sequences and genes that encode homologues of the new proteins. The  
XX present sequence encodes a portion of a soybean Arabidopsis thaliana-like  
XX sugar transport protein  
XX SQ Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. NO.: 2.32e-136 Length: 1692  
Score: 1670.50 Matches: 333  
Percent Similarity: 79.75% Conservative: 57  
Best Local Similarity: 68.10% Mismatches: 84  
Query Match: 40.99% Indels: 15  
DB: Gaps: 8  
US-10-051-909-32 (1-800) x ABK51966 (1-1692)  
Qy 308 AspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgPro 327  
Db 21 GAAAAAGATCAAATTAAGTTGTATGGACCAGAAACAAGGCCAGTCTGGGTTGCTAGACCT 80  
Qy 328 SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347



Db 81 GTTGCTGGACCAAAATCTGTTGGC-----CTTGATCTAGGAAAGCAATG 128  
QY 348 ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu 367  
Db 129 GCAAATCCAAGC---AGTCTAGTGGACCCCTCTAGTGACCCTCTTTGGTAGGTACATGAG 185  
QY 368 AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387  
Db 186 AAGCTCCCAAGAAACAGGA-----AGCACCCCTTTTCCACACTTTGGGAGTATG 233  
QY 388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTirAspGluGluAsnLeuHisArg 407  
Db 234 TTCAGTGTGGGGAAATCAGCCCAAGGAATGAAGATTGGGATGAGGAAGCCTAGCCAGA 293  
QY 408 AspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer 427  
Db 294 GAGGGTGATGATATGTCTCTGAT-----GCTGGTGATTCTGATGACAATTTGCAGAGT 347  
QY 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis 447  
Db 348 CCATTGATCTCAGCTCAAAACAACGAGTCTGGAT---AAGGACATACCTCCTCATGCCCAT 404  
QY 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGluGlyGlyAspGly 466  
Db 405 AGTAACCTTTGCA---AGCATGAGGCAAGGTAGTCTTTTACATGGAATTCAGGAGAACC 461  
QY 467 ValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTyrLysTyrSerGluLysGlu 486  
Db 462 ACTGGTAGTACTGGGATTGGTGGTTGCGAGCTAGCATGGAAATGGTCTGAAAGAGAG 521  
QY 487 GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal 506  
Db 522 GGCCAGATGGAAAGAAGGAGGTGGCTTCAAGAGAAATATATTTACACCAAGATGGTGGT 581  
QY 507 ProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGly 526  
Db 582 TCTGGATCTAGACGTGGGTCTGTGGTTTCACTCCCT---GGCGGTGATTACCAACTGAC 638  
QY 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546  
Db 639 AGTGAGGTTGTACAGGCTGCTGCTCTGGTGAGTCAGCCTGCCCTTATATAATGAGGACCTT 698  
QY 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566  
Db 699 ATCGCTCAACGCCCAGTTGGACCAGCTATGATTCATCCCTCTGAACAATTGCAAAAGGG 758  
QY 567 SerArgTyrLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGly 586  
Db 759 CCAAGTTGGAGTGATCTTTTGAACCTTGGGTGAAGCATGCAATTGATGTGGGGTGGGA 818  
QY 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle 606  
Db 819 ATGCAAAATCTTCAGCAGATTCTCTGGTATAAATGGGGTCCCTCTACTATACGCCTCAAATT 878  
QY 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626  
Db 879 CTTGAGCAGGCAGGTGTTGGTTATCTTCTTCAAGCCTAGGCCTTGGTTCTACTTCTTCA 938  
QY 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646  
Db 939 TCCTTTCTTATTAGTCGGGTGACAACCTTGTGATGCTTCTCTTGTATAGCCATTGCCATG 998  
QY 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666  
Db 999 AGGCTCATGGATATTTCAGGCAGAAGGACTTGTGCTCAGTACATAATCCCGTCTCTAATA 1058  
QY 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686  
Db 1059 GCAGCTCTTCTCATATTAGTCTCTGGGAAGTCTGTGGATTGTGGATTCACACTGCAAAATGCA 1118  
QY 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706

Db 1119 TCAATCTCAACCATTAGTGTATTGTCTATTCTGTTTCTTTGTCTATGGGATTGGACCA 1178  
QY 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726  
Db 1179 ATTCTTAATATACTTTGTGCAGAGATCTTCCCACACTCGAGTTCGTGGTCTCTGCATTGCT 1238  
QY 727 IleCysAlaPheThrPheTyrPileGlyAspIleIleValThrTyrSerLeuProValMet 746  
Db 1239 ATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGTCACTACACACTCCCAGTTATG 1298  
QY 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766  
Db 1299 CTCAATTCTGTAGGCCCTCGCTGGTGTTTTGGTATTATGCTGTCGTGCTTCATAGCA 1358  
QY 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786  
Db 1359 TGGGTGTTTGTCTTTTGAAGTTCAGAAACCAAGGGCATGCCACTGGAAGTGATCATT 1418  
QY 787 GluPhePheAlaValGlyAlaLysGln 795  
Db 1419 GAGTTCCTCTCTGTCGGAGCAAAACAG 1445  
RESULT 13  
ABX93202  
ID ABX93202 standard; cDNA; 1692 BP.  
XX  
AC ABX93202;  
XX  
DT 29-MAY-2003 (first entry)  
XX  
DE cDNA encoding soybean sugar transport protein #2.  
XX  
KW Arabidopsis thaliana-like sugar transport protein; corn; rice; wheat;  
KW plant sugar transport protein; carbohydrate transport; soybean;  
KW carbohydrate distribution; plant; gene; ss.  
XX  
OS Glycine max.  
XX  
PN US2002178468-A1.  
XX  
PD 28-NOV-2002.  
XX  
PF 17-JAN-2002; 2002US-00051902.  
XX  
PR 24-APR-1998; 98US-0083044P.  
PR 14-APR-1999; 99US-00291922.  
XX  
PA (ALLE/) ALLEN S M.  
PA (HITZ/) HITZ W D.  
PA (KINN/) KINNEY A J.  
PA (TING/) TINGEY S V.  
XX  
PI Allen SM, Hitz WD, Kinney AJ, Tingey SV;  
XX  
DR WPI; 2003-340957/32.  
DR P-PSDB; ABU08330.  
XX  
PT Novel plant sugar transport proteins and nucleic acid encoding the  
PT protein useful for producing transgenic plants having altered levels of  
PT sugar transport protein.  
XX  
PS Claim 2; Page 22-23; 56pp; English.  
XX  
CC The present invention relates to the isolation of Arabidopsis thaliana-  
CC like or Beta vulgaris-like sugar transport proteins, and the  
CC polynucleotide sequences encoding them. The plant sugar transport  
CC proteins of the invention have been isolated from corn, rice, soybean,  
CC and wheat. The polypeptides of the invention may be used for altering the  
CC level of expression of a sugar transport protein in a host cell, by  
CC transforming a host cell with a chimeric construct encoding all, or a  
CC portion of the sugar transport protein, in sense or antisense  
CC orientation. Particularly, the polypeptides may provide a means to  
CC control carbohydrate transport and distribution in plants. ABX93198-

CC	ABX93205 represent cDNA sequences encoding Arabidopsis thaliana-like									
CC	sugar transport proteins									
XX										
SQ	Sequence 1692 BP; 430 A; 339 C; 400 G; 523 T; 0 U; 0 Other;									
Alignment Scores:										
Pred. No.:	2.32e-136	Length:	1692							
Score:	1670.50	Matches:	333							
Percent Similarity:	79.75%	Conservative:	57							
Best Local Similarity:	68.10%	Mismatches:	84							
Query Match:	40.99%	Indels:	15							
DB:	8	Gaps:	8							
US-10-051-909-32 (1-800) x ABX93202 (1-1692)										
QY	308	AspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgPro	327							
DB	21	GAATAAGATCAAATTAAGTTGTATGGACCAGAACAAAGCCAGTCCTGGGTGCTAGACCT	80							
QY	328	SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet	347							
DB	81	GTTGCTGGACCAAAATTCGTGGC-----CCTGTATCTAGGAAAGGAAGCATG	128							
QY	348	ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu	367							
DB	129	GCAAAATCCAAAGC--AGTCTAGTGGACCCCTCTAGTGACCCCTCTTTGGTAGTGTACATGAG	185							
QY	368	AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet	387							
DB	186	AAGCTCCAGAAACAGGA-----AGCACCCCTTTTCCACACTTTGGGAGTATG	233							
QY	388	PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg	407							
DB	234	TTCAGTGTGGGGAAATCAGCCAAGGAATGAAGATGGGATGAGGAAAGCCTAGCCAGA	293							
QY	408	AspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer	427							
DB	294	GAGGGTGATGATTATGTCTCTGAT-----GCTGGTGATTCTGATGACAAATTGCAGAGT	347							
QY	428	ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis	447							
DB	348	CCATTGATCTCACGTCAAACACAGAGTCTGGAT---AAGGACATACCTCCTCATGCCCAT	404							
QY	448	ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGluGlyGlyAspGly	466							
DB	405	AGTAACCTTGCA--AGCATGAGGCAAGGTAGTCTTTTACATGGAAATTCAGGAGAAACCC	461							
QY	467	ValSerSerThrAspIleGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGlu	486							
DB	462	ACTGGTAGTACTGGGATTGGTGGTGGTGGCAGCTAGCATGGAAATGGTCTGAAAGAGAG	521							
QY	487	GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal	506							
DB	522	GGCCCAGATGGAAGAAGGAAGGTGGCTTCAAGAGAATATATTTACACCAAGATGGTGGT	581							
QY	507	ProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGly	526							
DB	582	TCTGGATCTAGACGTGGGTCTGTGGTTTCACTCCCT---GGCGGTGATTTACCAACTGAC	638							
QY	527	SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu	546							
DB	639	AGTGAGGTTGTACAGGCTGCTGCTCTGGTGAGTCAGCCCTGCCCTTTATATAGGACCTT	698							
QY	547	AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly	566							
DB	699	ATCGGTCAACGGCCAGTTGGACCAGCTATGATTCATCCCTCTGAAACAATTCGAAAAAGGG	758							
QY	567	SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGly	586							
DB	759	CCAAGTTGGAGTGATCTTTTGAACCTGGGGTGAAGCATGTCATGTGATGTGGGGGTGGGA	818							
QY	587	IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle	606							

819	ATGCAAAATCTTTCAGCAGTTCCTCTGGTATAAAATGGGGTCCTCTACTATACGCCTCAAATT	878
607	LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla	626
879	CTTGAGCAGGCAGGTGTTGGTTATCTTCTTCAAGCCTAGGCCCTTGGTTCTACTTCTTCA	938
627	SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet	646
939	TCCTTTCTTATTAGTCCGGTGACAACCTTGTGTGATGCTTCCTTGTATAGCCATTGCCATG	998
647	LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle	666
999	AGGCTCATGGATATTTTCAGGCAGAGGACTTTTGCTGCTCAGTACAATCCCCGTCCTAATA	1058
667	AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla	686
1059	GCAGCTCTTCTCATATTAGTCCTGGGAAGTCTTGTGGATTTGGGATCCACTGCAAAATGCA	1118
687	LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro	706
1119	TCAATCTCAACCATTAGTGTATTGTCTATTTCGTCTTTCTTTGTCATGGGATTTGGACCA	1178
707	IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla	726
1179	ATTCCTTAATATACTTTGTGCAGAGATCTTCCCACCTCGAGTTCGTGGTCTCTGCATTGCT	1238
727	IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet	746
1239	ATTTGTGCCCTTACCTTTTGGATCTGTGATATCATTTGTACCTACACACTCCCAGTTATG	1298
747	LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer	766
1299	CTCAATTCTGAGGCCTCGCTGGTGTTTTTGGTATTTAATGCTGCTCGTGTGCTTCATAGCA	1358
767	PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr	786
1359	TGGGTGTTTGTCTTTTGGAAAGTTCCAGAAACCAAGGGCATGCCACTGGAAGTGATCATT	1418
787	GluPhePheAlaValGlyAlaLysGln	795
1419	GAGTTCTTCTCTGTGCGAGCAAAACAG	1445
RESULT 14		
ADG47913		
ID	ADG47913 standard; cDNA; 1692 BP.	
XX		
AC	ADG47913;	
XX		
DT	11-MAR-2004 (first entry)	
XX		
DE	Soybean Arabidopsis-like sugar transport protein cDNA.	
XX		
KW	Arabidopsis-like sugar transport protein;	
KW	Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;	
KW	Soybean; plant; gene; ss.	
XX		
OS	Glycine max.	
XX		
FH	Location/Qualifiers	
FT	9..1469	
FT	/*tag= a	
FT	/product= "Soybean Arabidopsis-like sugar transport	
FT	protein "	
FT	/note= "CDS does not include start codon"	
FT	/partial	
XX		
PN	US2002199217-A1.	
XX		
PD	26-DEC-2002.	
XX		
PF	17-JAN-2002; 2002US-00051909.	
XX		
PR	24-APR-1998; 98US-0083044P.	
PR		





```
KW carbohydrate transport; grain filling; annual field crop; plant; gene;
KW ss.
XX
OS Triticum aestivum.
XX
FH Key Location/Qualifiers
FT CDS 3..1040
FT /*tag= a
FT /partial
FT /product= "portion of a wheat Arabidopsis thaliana-like
FT sugar transport protein"
FT /note= "This sequence lacks a start codon"
XX
XX
PN US6383776-B1.
XX
PD 07-MAY-2002.
XX
XX 14-APR-1999; 99US-00291922.
XX
XX 24-APR-1998; 98US-0083044P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Allen SM, Hitz WD, Kinney AJ, Tingey SV;
PI
XX
XX WPI; 2002-453364/48.
DR P-PSDB; AAU97207.
DR
XX
PT New nucleic acid encoding plant sugar-transport proteins, useful for
PT preparing transgenic plants with altered carbohydrate distribution.
PT
XX
PS Claim 3; Col 47-50; 54pp; English.
XX
XX The present invention relates to the isolation of plant polynucleotide
CC sequences encoding an Arabidopsis thaliana-like sugar transport protein
CC or Beta vulgaris-like sugar transport protein. The polynucleotide
CC sequences are useful for altering the level of sugar transport proteins
CC in plants, i.e. for control of carbohydrate transport and distribution in
CC plant cells, e.g. during grain filling of annual field crops (e.g. corn,
CC rice, soybeans, and wheat), and, for studying carbohydrate flows and
CC sugar transport. The polynucleotide sequences can also be used to isolate
CC cDNA sequences and genes that encode homologues of the new proteins. The
CC present sequence representing a contig assembled from various wheat cDNA
CC clones encodes a portion of an Arabidopsis thaliana-like sugar transport
CC protein
XX
XX Sequence 1487 BP; 355 A; 320 C; 373 G; 439 T; 0 U; 0 Other;
SQ
```

Alignment Scores:

Pred. No.:	8.24e-115	Length:	1487
Score:	1424.00	Matches:	281
Percent Similarity:	89.97%	Conservative:	24
Best Local Similarity:	82.89%	Mismatches:	32
Query Match:	34.94%	Indels:	2
DB:	6	Gaps:	2

US-10-051-909-32 (1-800) x ABK51968 (1-1487)

```
QY 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLys 481
Db 12 GAGGGTGGGAGGCGAGTCAGCAGCAGCTGGTATTGGTGGGGGGTGGCAACTCGCATGGAAA 71
QY 482 TrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeu 501
Db 72 TGGTCGGAGCGACAGCGGAGGATGGCAAGAAAGGAGGCTTCAAAAGAAATCTACTTG 131
QY 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521
Db 132 CACCAAGAGGGGGTGGCCGACTCAAGAAAGGGGCTCTGTTGTTTCACTTCTCTGGTGGGGT 191
QY 522 AspValPheGlu---GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540
Db 192 GATGCCACGCAAGGGGGCAGTGGGTTTATACATGCTGCTGCTTTGGTAAGCCACTCGGCT 251
```

```
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMet---SerAspAlaAlaMetValHisPro 559
Db 252 CTTTACTCCAAGGATCTTATGGAAGAGCGTATGGCGCGTATGGCGCGTATGATTCATCCA 311
QY 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579
Db 312 TTGGAGGCGAGCTCCCAAAGGTTCAATCTGGAAGATCTGTTTGAACCTGGTGTGAGGCGT 371
QY 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyVal 599
Db 372 GCATTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 431
QY 600 LeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPhe 619
Db 432 CTCTACTATACTCCTCAAAATTCTGGAGCAAGCTGGTGTGGCTGTCTTCTTCCAATCTT 491
QY 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerSerLeuThrLeuLeuMetLeu 639
Db 492 GGCCTCAGTTCAGCATCAGCATCCATCTTGATCAGTTCCTCTCACCACCTTACTCATGCTC 551
QY 640 ProCysIleGlyPheAlaMetLeuLeuMetLeuMetAspLeuSerGlyArgArgPheLeuLeu 659
Db 552 CCAAGCATTTGGTGTAGCCATGAGACTTATGGATATATCTGGAAGAAGGTTTCTGTACTG 611
QY 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAsp 679
Db 612 GGCACAAATTCCTATCTGATAGCATCCCTAATTGTTTGGGTGGTGGTCAATGTTATCAAC 671
QY 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699
Db 672 TTGAGTACGGTGGCCCGACGCTGTGCTCTCCACAGTTAGCGTCAATGTCTACTTCTGCTGC 731
QY 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719
Db 732 TTTGTCATGGGCTTTGGCCCGATCCCCAACATTTATGTGCAGAGATTTTCCCCACCAGA 791
QY 720 ValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739
Db 792 GTCCGTGGTGTCTGCACTCGCTATTTGGCCCTCACATTTCTGGATTTGTGACATTATTGTT 851
QY 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759
Db 852 ACCTACAGCCTGCCTGTGATGCTGAATGCTATTGGTCTAGCGGGTGTCTTTGGTATATAT 911
QY 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779
Db 912 GCAGTCGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
QY 780 MetProLeuGluValIleThrGluPhePheAlaValIcIlyAlaLysGlnAlaAlaAala 798
Db 972 ATGCCCTCGAGGTCATCACCGAGTCTTTCGCGGTTGGGGCGAAAGCAAGCGCAGGCC 1028
```

Search completed: April 13, 2005, 15:30:34  
Job time : 984.577 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 13:36:34 ; Search time 322.838 Seconds  
(without alignments)  
4054.730 Million cell updates/sec

Title: US-10-051-909-32  
Perfect score: 4075  
Sequence: 1 IRSGSWLAVQTPTPDLRR.....PLEVITEFFAVGAKQAAGA 800

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074038\_14080/app\_query.fasta\_1.1678  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909 @CGN\_1\_1\_141 @runat\_13042005\_074038\_14080 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3517	86.3	2824	3 US-09-291-922-1	Sequence 1, Appli
2	2674	65.6	2601	3 US-09-291-922-7	Sequence 7, Appli
3	1670.5	41.0	1692	3 US-09-291-922-9	Sequence 9, Appli
4	1424	34.9	1487	3 US-09-291-922-13	Sequence 13, Appl
5	943	23.1	1009	3 US-09-291-922-15	Sequence 15, Appli
6	562	13.8	870	3 US-09-291-922-5	Sequence 5, Appli
7	496	12.2	1853	3 US-09-291-922-23	Sequence 23, Appl
8	494.5	12.1	2017	3 US-09-291-922-21	Sequence 21, Appl
9	491.5	12.1	1914	3 US-09-291-922-19	Sequence 19, Appl
10	489	12.0	1872	3 US-09-291-922-27	Sequence 27, Appl
11	463.5	11.4	2089	3 US-09-291-922-25	Sequence 25, Appl
12	447	11.0	1431	4 US-09-489-039A-4762	Sequence 4762, Ap

13	440.5	10.8	1545	4 US-09-489-039A-4731	Sequence 4731, Ap
14	427	10.5	1506	4 US-09-489-039A-4560	Sequence 4560, Ap
15	426	10.5	1566	4 US-09-489-039A-2378	Sequence 2378, Ap
16	418.5	10.3	1668	4 US-09-614-221A-420	Sequence 420, App
17	412	10.1	4411529	3 US-09-103-840A-1	Sequence 1, Appli
18	368	9.0	4403765	3 US-09-103-840A-2	Sequence 2, Appli
19	355	8.7	2856	3 US-09-643-597-135	Sequence 135, App
20	355	8.7	2856	4 US-09-480-884A-135	Sequence 135, App
21	355	8.7	2856	4 US-09-542-615A-135	Sequence 135, App
22	355	8.7	2856	4 US-09-606-421B-135	Sequence 135, App
23	355	8.7	2856	4 US-09-221-107-135	Sequence 135, App
24	355	8.7	2856	4 US-09-466-396A-135	Sequence 135, App
25	355	8.7	2856	4 US-09-476-496A-135	Sequence 135, App
26	355	8.7	2856	4 US-09-630-940B-135	Sequence 135, App
27	355	8.7	2856	4 US-09-285-479-135	Sequence 135, App
28	354.5	8.7	3000	2 US-08-928-692-9	Sequence 9, Appli
29	354.5	8.7	3000	3 US-09-339-972-9	Sequence 9, Appli
30	352.5	8.7	1626	4 US-09-614-221A-521	Sequence 521, App
31	346	8.5	5228	4 US-09-919-039-216	Sequence 216, App
32	343	8.4	1776	4 US-09-679-686B-11	Sequence 11, Appl
33	338.5	8.3	1943	4 US-09-774-528-168	Sequence 168, App
34	334.5	8.2	5227	4 US-09-919-172-79	Sequence 79, Appl
35	332.5	8.2	2592	3 US-09-591-025-8	Sequence 8, Appli
36	328.5	8.1	2592	4 US-09-894-927B-8	Sequence 8, Appli
37	322	7.9	1659	4 US-09-248-796A-6610	Sequence 6610, Ap
39	312	7.7	1675	4 US-09-614-221A-534	Sequence 534, App
40	312	7.7	1695	4 US-09-614-221A-443	Sequence 1, Appli
41	307.5	7.5	1704	4 US-09-614-221A-96	Sequence 443, App
42	302	7.4	2298	4 US-09-248-796A-5965	Sequence 96, Appl
43	301.5	7.4	3915	4 US-09-023-655-1104	Sequence 5965, Ap
44	293	7.2	987	4 US-09-710-279-1941	Sequence 1104, Ap
45	293	7.2	3065	4 US-09-710-279-3697	Sequence 1941, Ap
					Sequence 3697, Ap

ALIGNMENTS

RESULT 1  
US-09-291-922-1  
; Sequence 1, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 1  
; LENGTH: 2824  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (29)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (622)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (636)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (638)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (669)

/	FEATURE:																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
---	----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Qy	274	SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle	293						
Db	898	GCTCTTCTAGTTGAAGTTTGGGGTTCGGTAAAGATACACGTATTTNAGAGTACATCATT	957						
Qy	294	GlyProAlaThrGluAlaAAspAspLeuValThrAspGlyAspLysGluGlnIleThr	313						
Db	958	GGACCTGCCACGAGGCAGCCGATGATCTTTGTAACGTGACGGTGATAAGGAACAATCACA	1017						
Qy	314	LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet	333						
Db	1018	CTTTATGGGCCTGAAGAAGGCCAGTCATGGATTGCTCGACCTTCTAAGGGACCCATCATG	1077						
Qy	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353						
Db	1078	CTTGGAAGTGCTTCTCTTGTCATCTCGTCATGGGAGCATGGTGAACCCAGAGTGTACCC	1137						
Qy	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373						
Db	1138	CTTATGGATCCGATTGTGACACTTTTGTGTAGTGTCCATGAGAAATATGCCTCAAGCTGGA	1197						
Qy	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393						
Db	1198	GGAAATATGAGGAGCACATTGTTTCCAAACTTTGGAAGTATGTTCAGTGTACAGATCAG	1257						
Qy	394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413						
Db	1258	CATGCCAAAATGAGCAGTGGGATGAAGAGAAATCTTCATAGGGATGACGAGGATACGCA	1317						
Qy	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433						
Db	1318	TCTGATGGTGCAGAGGTGACTATGAGGACAAATCTCCATAGCCCATTTGCTGCCAGGCAG	1377						
Qy	434	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer	453						
Db	1378	GCAACAGGTGCGGAAGGAAGGACATTGTGCACCATTGTCCCGTGGAAAGTGTCTTGAGC	1437						
Qy	454	MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly	473						
Db	1438	ATGAGAAGGCCAAAGCCTCTTAGGGAGGGTGGAGATGGTGTGACGACACTGATATCGGT	1497						
Qy	474	GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu	493						
Db	1498	GGGGGATGGCAGCTTGTCTGGAAATGGTCAGAGAAGGAAGGTGAGAATGGTAGAAGGAA	1557						
Qy	494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer	513						
Db	1558	GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGGAGTTCCTGGCTCAAGAAGGGCTCA	1617						
Qy	514	IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533						
Db	1618	ATTGTTTCACTTCCCGGTGGTGGCGATGTTCTTGAGGGTAGTGAAGTTGTACATGCTGCT	1677						
Qy	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553						
Db	1678	GCTTTAGTAAGTCAGTCAGCACACTTTTCTCAAAGGGTCTTGTGTAACCCACGCATGT	1737						
Qy	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe	573						
Db	1738	GCTGCCATGGTTTCAACCATCTGAGGTAGCTGCCAAAGGTTTCCAGTTGGAAAGATTTGTTT	1797						
Qy	574	GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe	593						
Db	1798	GAACTGGAGTGAGGCGTGCCCTGTTAGTCGGTGTGTGGAATTCAGATCTTCAACAGTTT	1857						
Qy	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla	613						
Db	1858	GCTGGAATAAACGGTGTCTGTACTATACCCCAAAATTTCTTGAGCAAGCTGGTGTGGCA	1917						
Qy	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeu	633						
Db	1918	GTTATTTCTTTCCAAATTTGGTCTCTCAGCTCGGCATCAGCATCCATCTTTGATCAGTTCTCTC	1977						





Db 1297 ACAACAAGCCTTGAA--AAAGACTTACCTCCTCCTCCTTCCCATGGCAGTATCCTTGGC 1353  
Qy 453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471  
Db 1354 AGCATGAGGCGTCACAGTAGTCTCATGCAAGGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413  
Qy 472 IleGlyGlyGlyTyrGlnLeuAlaTyrPlysTrpSerGluLysGluGlyGluAsnGlyArg 491  
Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAATGGACTGATAAA---GGTGAGGATGGAAAA 1470  
Qy 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511  
Db 1471 CAACAAGGAGGGTTTAAAGGATTTATTTACATGAGGAGGGAGTTTCTGCATCTCGTCGT 1530  
Qy 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531  
Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTGTCCAG 1572  
Qy 532 AlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551  
Db 1573 GCTGCTGCCTTGGTAAGCCAACCCGCTCTTTACTCCAAGGAGCTTATGTATGGACACCCA 1632  
Qy 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAsp 571  
Db 1633 GTTGGGCGCTGCAATGGTTACCCCATCTGAGACAGCTTCAAAGGGGCAAGTTGGAAAGCT 1692  
Qy 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591  
Db 1693 CTTCTTGAACCAAGGGTTAAGCATGCATTGGTTGTTGGAGTTGGAATACAAATACTTTCAG 1752  
Qy 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly 611  
Db 1753 CAGTTTTCAGGGATAAATGGGGTTCTATATTACACACCTCAAATCCTTGAAGAGGCCGGT 1812  
Qy 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631  
Db 1813 GTTGAAGTTCTTCTTTTCAGATATAGGCATTGGCTCAGAGTCGGGCATCATTCCTTATCAGT 1872  
Qy 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651  
Db 1873 GCTTTCACAACCTTCTTGATGCTTCCCTGTATAGGCGTAGCCATGAAGCTCATGGATGTT 1932  
Qy 652 SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671  
Db 1933 TCAGGCAGAAAGCGAGTTGCTACTTACTACAAATCCCCGTGCTGATTGTGTCACTCATTTAT 1992  
Qy 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691  
Db 1993 TTGGTCATTGGAAGCCTGGTAAATTTTGGCAATGTGCCCCATGCAGCAATCTCAACAGTA 2052  
Qy 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711  
Db 2053 TGGCTGTGTGTTTATTTCTGCTGCTTGTGTATGGGTATATGGACCAATTCAAAACATCCTT 2112  
Qy 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731  
Db 2113 TGCTCAGAGATTTTCCCCACTAGGGTGGCTGCTGCAATTGCTATCTGTGCATTAGTG 2172  
Qy 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751  
Db 2173 TTTCTGGATTGGAGACATCATCATCATACTCGCTGCGCTGTGATGCTCGGCTCTTTTAGGA 2232  
Qy 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771  
Db 2233 CTTGGTGGTGTATTTCGCCATTTACGCAGTTGTGTGTTTTCATCTCGTGGATATTTGTGTTT 2292  
Qy 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791  
Db 2293 TTGAAGGTTCCAGAAACAAGGGCATGCCCCCTTGAAGTCATCTCTGAATTCCTTTCTGTT 2352  
Qy 792 GlyAlaLysGlnAlaAlaAala 798

Db 2353 GGAGCAAAAGCAGGCTGCTTCT 2373  
RESULT 3  
US-09-291-922-9  
; Sequence 9, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 9  
; LENGTH: 1692  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-291-922-9  
Alignment Scores: 1.11e-166 Length: 1692  
Pred. No.: 1670.50 Matches: 333  
Score: 79.75% Conservative: 57  
Percent Similarity: 68.10% Mismatches: 84  
Best Local Similarity: 40.99% Indels: 15  
Query Match: 3 Gaps: 8  
DB:  
US-10-051-909-32 (1-800) x US-09-291-922-9 (1-1692)  
Qy 308 AspLysGluGlnIleThrLeuTyrGlyProGluGlyGlnSerTyrIleAlaArgPro 327  
Db 21 GAAAAAGATCAAATTAAGTTGTATGGACCAGAACAAAGGCCAGTCCTGGTTGCTAGACCT 80  
Qy 328 SerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMet 347  
Db 81 GTTGCTGGACCAAATTCGTITGGC-----CTTGATCTAGGAAAGGAAGCATG 128  
Qy 348 ValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGlu 367  
Db 129 GCAAATCCAAAGC--AGTCTAGTGGACCTCTAGTGACCCCTCTTGGTAGTGTACATGAG 185  
Qy 368 AsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMet 387  
Db 186 AAGCTCCCAGAAACAGGA-----AGCACCCCTTTTCCACACTTTTGGGAGTATG 233  
Qy 388 PheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArg 407  
Db 234 TTCAGTTTGGGGGAAATCAGCCCAAGGAATGAAGATTGGGATGAGGAAAGCCTAGCCAGA 293  
Qy 408 AspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSer 427  
Db 294 GAGGGTGTATGATTATGTCTCTGAT-----GCTGGTGAATTCGTATGACAAATTTGCAGAGT 347  
Qy 428 ProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHis 447  
Db 348 CCATTGATCTCAGTCAAACAACCGAGTCTGGAT--AAGGACATACCTCCTCATGCCCAT 404  
Qy 448 ArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu---GlyGluGlyGlyAspGly 466  
Db 405 AGTAACCTTGCA--AGCATGAGGCAAGGTAGTCTTTTACATGGAAATTCAGGAGAACCC 461  
Qy 467 ValSerSerThrAspIleGlyGlyTyrGlnLeuAlaTyrLysTrpSerGluLysGlu 486  
Db 462 ACTGGTAGTACTGGGATTGGTGGTGGTGGCAGTAGCATGGAAATGGTCTGAAAGAGAG 521  
Qy 487 GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal 506

Db 522 GSCCCAGATGGAAGAAGGAGGTGGCTTCAAGAGAAATATATTACACCAAGATGGTGGT 581

Qy 507 ProGlySerArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGly 526

Db 582 TCTGGATCTAGACGTGGTCTGTGGTTTCACCTCCCT--GGCGGTGATTACCAACTGAC 638

Qy 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546

Db 639 AGTGAGTTGTACAGGCTGCTGCTGGTGAGTCAGCCTGCCCTTATATAATGAGGACCTT 698

Qy 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566

Db 699 ATGCGTCAACGGCCAGTTGGACCAGCTATGATTTCATCCCTCTGAAACAATTGCCAAAGGG 758

Qy 567 SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuValGlyValGly 586

Db 759 CCAAGTTGGAGTGATCTTTTGAACCTGGGGTGAAGCATGCATTGATTGTGGGGTGGGA 818

Qy 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle 606

Db 819 ATGCAAAATCTTCAGCAGTTCTCTGTTATAAATGGGGTCCCTCTACTATACGCCTCAAAAT 878

Qy 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626

Db 879 CTTGAGCAGCAGGTGTGGTTATCTCTTTCAAGCCTAGGCCCTGGTTCCTACTTCTTCA 938

Qy 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646

Db 939 TCCTTTCTTATTAGTCGGTGACAACTTGTTGATGCTTCCTTGTATAGCCATTGGCCATG 998

Qy 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIle 666

Db 999 AGGCTCATGGATATTTTCAGGCAGAGGACTTTGCTGCTCAGTACAAATCCCGTCTCTAATA 1058

Qy 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686

Db 1059 GCAGCTTCTCATATTAGTCCCTGGGAAGTCTTGTGGATTTGGGATCCACTGCAAATGCA 1118

Qy 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706

Db 1119 TCAATCTCAACCATTAGTGTATTGTCTATTCTGTTCTTTGTCTCATGGGATTTGGACCA 1178

Qy 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726

Db 1179 ATTCCTAATATACTTTGTGCAGAGATCTTCCCCACTCGAGTTCGTGCTCTGCAATTGCT 1238

Qy 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746

Db 1239 ATTTGTGCCCTTACCCTTTTGGATCTGTGATATCATTTGTACCTTACACACTCCCAGTTATG 1298

Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer 766

Db 1299 CTCAATTCGTAGGCCTCGCTGGTGTGTTTGGTATTTATGCTGTCGTGCTTCATAGCA 1358

Qy 767 PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr 786

Db 1359 TGGGTGTTTGTCTTTTGAAGTTCCAGAAACCAAGGGCATGCCACTGCCAGTGATCATT 1418

Qy 787 GluPhePheAlaValGlyAlaLysGln 795

Db 1419 GAGTTCCTCTGCTCGGAGCAAAACAG 1445

RESULT 4

US-09-291-922-13

; Sequence 13, Application US/09291922

; Patent No. 6383776

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: BB-1163

; CURRENT APPLICATION NUMBER: US/09/291,922

; CURRENT FILING DATE: 1999-04-14

; EARLIER APPLICATION NUMBER: 60/083,044

; EARLIER FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 13

; LENGTH: 1487

; TYPE: DNA

; ORGANISM: Triticum aestivum

US-09-291-922-13

Alignment Scores:

Pred. No.: 1.27e-140 Length: 1487

Score: 1424.00 Matches: 281

Percent Similarity: 89.97% Conservative: 24

Best Local Similarity: 82.89% Mismatches: 32

Query Match: 34.94% Indels: 2

DB: 3 Gaps: 2

US-10-051-909-32 (1-800) x US-09-291-922-13 (1-1487)

Qy 462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLys 481

Db 12 GAGGGTGGGAGGCAGTTCAGCAGCACTGGTATTGTTGGGGGGTGGCAACTCGCATGGAAA 71

Qy 482 TrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeu 501

Db 72 TGGTCGAGCGACAAGGCGAGGATGGCAAGAAAGGAGGCTTCAAAAGAATCTACTTG 131

Qy 502 HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGly 521

Db 132 CACCAAGAGGGGGTGGCCGACTCAAGAAAGGGGCTCTGTTGTTTCACTTCCTTGGTGGGGT 191

Qy 522 AspValPheGlu--GlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540

Db 192 GATGCCACGCAAGGGGGCAGTGGGTTTATACATGCTGCTGCTTTGGTAAGCCACTCGGCT 251

Qy 541 LeuPheSerLysGlyLeuAlaGluProArgMet--SerAspAlaAlaMetValHisPro 559

Db 252 CTTTACTCCAAGGATCTTTATGGAAGAGCGTATGGCGGCCGTCAGCCATGATTCATCCA 311

Qy 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579

Db 312 TTGGAGGCAGCTCCCAAAGGTTCAATCTGGAAAGATCTGTTTGAACCTGGTGTGAGGCGT 371

Qy 580 AlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyVal 599

Db 372 GCATTGTTCTGTCGGTGTGGAATTCAGATGCTTCAGCAGTTTGTCTGGAATAAATGGAGTT 431

Qy 600 LeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPhe 619

Db 432 CTCTACTATACTCCTCAAAATCTGGAGCAAGCTGGTGTGGCTGTCTTCTTTCCAATCTT 491

Qy 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639

Db 492 GGCCTCAGTTCAGCATCAGCATCCATCTTGATCAGTTCTCTCACCACCTTACTCATGCTC 551

Qy 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeu 659

Db 552 CCAAGCATTTGGTGAGCCATGAGACTTATGGATATATCTGGAAGAAGGTTTCTGCTACTG 611

Qy 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAop 679

Db 612 GGCACAATTCCTCATCTTGATAGCATCCCTAAATTGTTTGGGTGTGGTCAATGTTATCAAC 671

Qy 680 LeuGlyThrLeuAlaHisAlaLeuSerThrValSerValIleValTyrPheCysCys 699

Db 672 TTGAGTACGGTGCCCCACGCTGTGCTCTCCACAGTTAGCGTCAATTGTCTACTTCTGTGTC 731

Qy 700 PheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArg 719

Db 732 TTTGTCTATGGGCTTTGGCCCCGATCCCCAACATTCATATGTGCAGAGATTTTCCCCACCAGA 791





Qy 749 AlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheVal 768  
|||||  
Db 248 GCCATTGGACTCGCTGGAGTGTGGAAATCTACGCAGTGGTCTGCATACTGGCTTTCCTG 307  
  
Qy 769 PheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhe 788  
|||||  
Db 308 TTTGTCTTCATGAAGGTGCCGGAGACAAAGGCATGCCTCTTGAAGTCATCACCGAGTTC 367  
  
Qy 789 PheAlaValGlyAlaLysGlnAla 796  
|||||  
Db 368 TTCTCTGTCGGAGCAAAGCAGGCC 391

RESULT 7  
US-09-291-922-23  
; Sequence 23, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; EARLIER FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 23  
; LENGTH: 1853  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-291-922-23

Alignment Scores:  
Pred. No.: 5.6e-42 Length: 1853  
Score: 496.00 Matches: 165  
Percent Similarity: 36.39% Conservative: 109  
Best Local Similarity: 21.91% Mismatches: 189  
Query Match: 12.17% Indels: 290  
DB: 3 Gaps: 18

US-10-051-909-32 (1-800) x US-09-291-922-23 (1-1853)

Qy 59 AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAla 78  
|||  
Db 257 GCTTTTGGTGTGTATGCTGGCCTCCATGACTTCCATCTTGGTGTATGATATTGGA 316  
  
Qy 79 ThrIleAlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGlu----- 96  
: : : : :  
Db 317 GTGATGAGTGGAGCAGCCATATACATAAAAGGACCTGAAAGTCTCGGACGACAAATC 376  
  
Qy 97 ProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThr 116  
: : : : :  
Db 377 GAGATCTGCTCGGAATCATCAACCTATACTCTCTGATAGGCTCATGTCTC----- 427  
  
Qy 117 PheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIle 136  
: : : : :  
Db 428 ---GCCGGCAGAACTCCGACTGGATAGGTCCCGTTACACGATTGTTTTCGCCGGCACC 484  
  
Qy 137 LeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuLeu 156  
: : : : :  
Db 485 ATCTTCTTGTGCGGAGCACTTCTCATGGGTTTCTCCCCCAATTATTCTCTCATGTTT 544  
  
Qy 157 AlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIle 176  
|||||  
Db 545 GGCCGTTTGTGCTGGCATTTGGCATCGGCTACGCCCTCATGATAGCCCCCGTCTACACC 604  
  
Qy 177 SerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGly 196  
: : : : :  
Db 605 GCCGAGGTCTCCCCGGCCTCCTCTCTGTGGCTTCCCTCACTTCTCCTTGAGGTATTATT 664

Qy 197 SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSer---LeuSerProSer 215  
: : : : :  
Db 665 AATGGAGGGATATTAAATTGGATACATATCAAACTATGCAATTTTCGAAGCTGACACTAAAG 724  
  
Qy 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
|||||  
Db 725 GTGGGATGGCGAATGATGCTTGGAGTTGGTGTCAATACCTTCGGTA---CTCCTAACACGTA 781  
  
Qy 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
: : : : :  
Db 782 GGAGTGTGGCGATGCCGGAGTCCCCAAGGTGGCTTGTGATGAGGGTTCGTTTGGGAGAG 841  
  
Qy 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
|||||  
Db 842 GCAAGAAAAGTGTCTTAACAAA----- 862  
  
Qy 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
Db 862 ----- 862  
  
Qy 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
|||  
Db 863 -----ACCTCAGACAGCAAGCAAGAG----- 883  
  
Qy 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
Db 883 ----- 883  
  
Qy 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355  
: : : : :  
Db 884 GCCCAACTAAGGCTAGCG----- 901  
  
Qy 356 AspProIleValThrLeupheGlySerValHisGluAsnMetProGlnAlaGlyGlySer 375  
: : : : :  
Db 902 -----GAATCAAAACAGCCGAGGGATC 925  
  
Qy 376 MetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAla 395  
|||  
Db 926 CCCGAGAGTTGC----- 937  
  
Qy 396 LysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAsp 415  
|||  
Db 938 -----AACGACGAC----- 946  
  
Qy 416 GlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThr 435  
Db 946 ----- 946  
  
Qy 436 GlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArg 455  
: : : : :  
Db 947 -----GTCGTTCAGGTAAT 961  
  
Qy 456 ArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly 475  
: : : : :  
Db 962 AAACAAAGC-----AACGGTGAAGGTGTA----- 985  
  
Qy 476 TrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGly 495  
Db 985 ----- 985  
  
Qy 496 PheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgGlySerIleVal 515  
Db 985 ----- 985  
  
Qy 516 SerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeu 535  
Db 985 ----- 985  
  
Qy 536 ValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAla 555  
Db 985 ----- 985

QY 556 MetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe----- 573  
Db 986 -----TGGAAAGAGCTCTTCCTCTAT 1006  
QY 574 ---GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGln 592  
Db 1007 CCAACGCCGCAATTCTGTACATCGTAATCGCTGCCCTTGGTATTCACTTCTTCCAACAA 1066  
QY 593 PheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyVal 612  
Db 1067 GCGTCGGGCGTAGACGCCGTGTTTGTACAGCCCGAGATCTTCGAAAAGGCTGGGATT 1126  
QY 613 AlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSer 632  
Db 1127 ACAACGACACGCATAAGCTT-----CTTGCAACCGTGGCCGTTGGATTTC 1171  
QY 633 LeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSer 652  
Db 1172 GTTAAGACCGTGTTCATCTTGGCG-----GCTACGTTTACGTTGGACCGCGTG 1219  
QY 653 GlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672  
Db 1220 GGTTCGTCTCCGTTGTTATGTCTAGTGTCCGGCGCATGGTCTCTCGTTTCTCACGCTT 1279  
QY 673 ValValSer---AsnLeuIleAspLeuGly-----ThrLeuAlaHisAlaLeuLeuSer 689  
Db 1280 GCGATCAGCCTCACTGTATTATGATCATTCGGAGAGGAAATTAATGTGGGCCGTTGGATCG 1339  
QY 690 ThrValSer---ValIleValTyrPheCysCysPheValMetGlyPheGlyProIlePro 708  
Db 1340 AGCATAGCCATGGTGTGGCTTACGTGGCCACGTTCTCCATCGGTGGGTCCCATCAG 1399  
QY 709 AsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCys 728  
Db 1400 TGGGTCTATAGTTCTGAGATCTTCCCGTTGAGGTGCGGGCGCARGGTGCGGCGGGA 1459  
QY 729 AlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsn 748  
Db 1460 GTTGGCGTGAATAGGACCACTAGCGCGGTGTGTCATATGATGATTTTCTGTCCCTCACTAGA 1519  
QY 749 AlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheVal 768  
Db 1520 GCCATCACTATTGGTGGAGCTTCTTCTCTTATTGTGGCATTGCTACTGTGGGTGGATA 1579  
QY 769 PheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr---Glu 787  
Db 1580 TTCTTTTACACCGTCTTGCTGTGAGACCCGGGGAACCGTCTCGAAGACATGGAAGGTCT 1639  
QY 788 PhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800  
Db 1640 TTTGGTACTTTTAGTCCAAATCCAACGCCAGCAAGGCT 1678

RESULT 8

US-09-291-922-21  
; Sequence 21, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 21  
; LENGTH: 2017  
; TYPE: DNA  
; ORGANISM: Oryza sativa

US-09-291-922-21  
Alignment Scores: 9.38e-42 Length: 2017  
Pred. No.: 494.50 Matches: 185  
Score: 35.71% Conservative: 101  
Percent Similarity: 23.10% Mismatches: 206  
Best Local Similarity: 12.13% Indels: 309  
Query Match: 3 Gaps: 22  
DB:  
US-10-051-909-32 (1-800) x US-09-291-922-21 (1-2017)  
QY 24 LeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCysSerSerGln 43  
Db 30 TTACACTCGACCGCCACTACTGTACACGGCCCA--GAGCGAGCCTCCTCCTCCTTGCA 86  
QY 44 GluProValThrSerAspAspIleLeuGluAsp-----LysMetSerGly--- 58  
Db 87 CCACCGGAGATGGCTTCGCCCGCGCTGCCGGAGGCCGTCGCGCGGAGGAGGCAAC 146  
QY 59 -----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75  
Db 147 GTCCGGTTCGCCTTCGCCTTCGCCCATCCTCGCCTCCATGACCTCCATCCTCCTCGGCTAC 206  
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95  
Db 207 GATATCGGGTGATGAGCGGGCGCTCGCTGTACATCAAGAAGGACTTCAACATC--AGT 263  
QY 96 GluProThrValGlu-----GlyLeuIleValSerMetSerLeuIleGlyAlaThr 112  
Db 264 GACGGGAAGTGGAGGTTCTCATGGGCATACTGAACCTCTACTCGCTCATCGGCTCCTTC 323  
QY 113 IleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIle 132  
Db 324 GCG-----GCGGGGCGGACGTGGACTGGATCGGCGCGGTACACCATCGTG 371  
QY 133 LeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyr 152  
Db 372 TTCGCCCGCGTCATATTCTTCGCGGGGSGGTTCCCTCATGGGGTTCGCGCTCAACTACGCC 431  
QY 153 ValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuVal 172  
Db 432 ATGCTCATGTTCCGCCGCTTCGTGGCCGCATCGCGGTGGGCTACGCGCTCATGATCGCG 491  
QY 173 ProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuPro 192  
Db 492 CCGGTGTACACCGCGAGGTGTCCGCGGCGTCCGCGCGTGGCTTCTGACGTCTCCCG 551  
QY 193 GlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeu 212  
Db 552 GAGGTGTTTCATCAACTTCGGCATCCTGCTCGGTACGTTCTCGAACTATGCTTCTCCCGC 611  
QY 213 SerPro---SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe 231  
Db 612 TTGCCGCTGAACCTCGGGTGGCGCATCATGCTCGGCATCGGCGCGGCGCGTCCGTG--- 668  
QY 232 PhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGly 251  
Db 669 CTGCTCGCGCTCATGGTGTCTCGGCATGCCGAGTCCGCCGCTGGTGGTGGTGGTGGTGG 728  
QY 252 ArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGly 271  
Db 729 CGCCTCGCGGACGCCAAGGTGGTGTGGAGAAGACC----- 764  
QY 272 GluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyr 291  
Db 765 -----TCCGACACG----- 773  
QY 292 IleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGln 311  
Db 774 -----GCGGAGGAGCGCGGAGCGCGCTG----- 797  
QY 312 IleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyPro 331



Db 797 ----- 797  
QY 332 IleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSer 351  
Db 797 ----- 797  
QY 352 ValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGln 371  
Db 797 ----- 797  
QY 372 AlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThr 391  
Db 797 ----- 797  
QY 392 AspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGlu 411  
Db 797 ----- 797  
QY 412 TyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSer 431  
Db 797 ----- 797  
QY 432 ArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAla 451  
Db 797 ----- 797  
QY 452 LeuSerMetArgArgGlnThrLeuLeuGlyGlyGluGlyAspGlyValSerSerThrAsp 471  
Db 798 -----GCCGAC 803  
QY 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491  
Db 804 ATCAAGCCGCC----- 815  
QY 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511  
Db 816 -----GCCGGCATCCTGAGGAGCTCGAC 839  
QY 512 GlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGluGlySerGluPhe 529  
Db 840 GCGACGTGTGACCGTCCCAAGAGAGGCGGAAC----- 878  
QY 530 ValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluPro 549  
Db 879 -----GAGAAG 884  
QY 550 ArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyr 569  
Db 885 CGGGTG-----TGG 893  
QY 570 LysAspLeuPhe-----GluProGlyValArgArgAlaLeuLeuValGlyValGly 586  
Db 894 AAGGAGCTCATCTGTCCCCGACCCCGCATGCGGGCATCTGTGTCGGGATCGGC 953  
QY 587 IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle 606  
Db 954 ATCCACTTCTCCAGCATGCGTTGGGCATTCACTCCGTCGTCTTCTACAGCCCTCTCGTG 1013  
QY 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626  
Db 1014 TTCAAGAGCCCGGATTACG-----AACGACAAACACTTCTTGGGCACCACTTGGCCG 1067  
QY 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646  
Db 1068 TTCGGTGTACCAAGAGGCTTTTCATCTGTG-----GCCACT 1106  
QY 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666  
Db 1107 TTCTTCATCGACGGCGTCCGGCGCGCGCTGTTGTGGCAGCACGGGGGGGATAATC 1166  
QY 667 AlaSerLeuVal-----IleLeuValValSerAsnLeuIleAspLeuGly 681

Db 1167 CTCTCCCTCATCGGCCCTCGGGCGCCGGGCTCACCGTCGTGCGGCCAGCACCCCGACGCCAAG 1226  
QY 682 ThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal 701  
Db 1227 ATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTCGCCTACGTGCGCTTCTTCTCC 1286  
QY 702 MetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArg 721  
Db 1287 ATCGGCCTTGGCCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCGCTCCAGGTGCGC 1346  
QY 722 -----GlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleVal 739  
Db 1347 GCGCTGGGCTGCTCGCTCGGGCTCGCGGCCAACCGCGTCAACCAGCGGCTCATCTCCATG 1406  
QY 740 ThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyr 759  
Db 1407 ACCTTCCTG-----TCGCTGTCCAAGGCCATCACCTACCTACCTCCCGGAGACCCGCGC 1460  
QY 760 AlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGly 779  
Db 1461 TCCGCGCATCGCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTCCCGGAGACCCGCGC 1520  
QY 780 MetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLys 799  
Db 1521 CGGACGCTGGAGGAGATGAGCAAGCTGTTC-----GGCGACACGGCGCGCTCGGAA 1574  
QY 800 Ala 800  
Db 1575 TCA 1577  
RESULT 9  
US-09-291-922-19  
; Sequence 19, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; EARLIER FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 19  
; LENGTH: 1914  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-09-291-922-19  
Alignment Scores:  
Pred. No.: 1.78e-41 Length: 1914  
Score: 491.50 Matches: 172  
Percent Similarity: 35.36% Conservative: 107  
Best Local Similarity: 21.80% Mismatches: 201  
Query Match: 12.06% Indels: 309  
DB: 3 Gaps: 18  
US-10-051-909-32 (1-800) x US-09-291-922-19 (1-1914)  
QY 41 SerSerGlnGluProValThr-----SerAspAspIle-----LeuGluAspLys 55  
Db 35 TCAAGAAGTAGCCGTTAACGATGGCTTCCGACGAGCTCGCAAAAGGCCGTCGAGCCCAGG 94  
QY 56 MetSerGly-----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeu 71  
Db 95 AAGAAGGCAACGTCAAAGTATGCTCCATATGTGCCATCTCTGGCTCCATGGCCTCTGTG 154  
QY 72 LeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPhe 91

Db 155 ATCCTTGGCTATGACATFTGGGTGATGAGTGGAGCGGCCATGTAATCAAGAAGGACCTG 214  
QY 92 GlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla 111  
Db 215 AATATC--ACGGACGTGCAGCTGGAGATCCTGATCGGGATCCTCAGTCTCTAC----- 265  
QY 112 ThrIleValThrThrPheSerGlyPro---LeuSerAspSerIleGlyArgArgProMet 130  
Db 266 TCGCTGTTTCGGATCCTTCGCTGGCGCGCGACGTCCGACAGGATCGGGCGCGCTTGACC 325  
QY 131 LeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsn 150  
Db 326 GTCGTGTTTCGCCGTGTCACTTCTTCGTGGGCTCGTTGCTCATGGGTTTCGCCGTCAAC 385  
QY 151 ValTyrValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThr 170  
Db 386 TACGGCATGCTCATGGCGGCGCGCTTCGTGGCCGAGTCGGTGTGGGTACGGGGGCAIG 445  
QY 171 LeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThr 190  
Db 446 ATCGCGCCCGTGTACACGGCGAGATCTCGCCTGCGCGCTCCCGTGGCTTCCTGACCACC 505  
QY 191 LeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMet 210  
Db 506 TTCCCGGAGGTGTTTCATCAACATCGGCATCCTGCTGGCTACCTGTCCAACTTCGCGTTC 565  
QY 211 SerLeuSerPro---SerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSer 229  
Db 566 GCGCGCCTCCCGCTCCACCTCGGCTGGCGCGTCACTGCTGCCATTCGGCAGTTCGGTCC 625  
QY 230 LeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSer 249  
Db 626 ---GGCCTGCTCGCGTCTCTGGTGTCTGCATGCCCGAGTCGCCTCGGTGGTCTTG 682  
QY 250 LysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspVal 269  
Db 683 AAGGGCCGCTCGCGACGCCAGCGCTGTGCTAGAGAAG----- 721  
QY 270 SerGlyGluLeuSerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGlu 289  
Db 721 ----- 721  
QY 290 GluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLys 309  
Db 721 ----- 721  
QY 310 GluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys 329  
Db 722 -----ACCTCTGCCACGCCAGAGAGCGCCCGAGCGCTGGCC----- 760  
QY 330 GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn 349  
Db 760 ----- 760  
QY 350 GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet 369  
Db 760 ----- 760  
QY 370 ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer 389  
Db 760 ----- 760  
QY 390 ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAsp 409  
Db 760 ----- 760  
QY 410 GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeu 429  
Db 760 ----- 760  
QY 430 LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGly 449  
Db 760 ----- 760

QY 450 SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSer 469  
Db 760 ----- 760  
QY 470 ThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsn 489  
Db 761 ---GACATCAAGCCCGCG----- 775  
QY 490 GlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer 509  
Db 776 -----GCGGGGATTCCGAAGGGC 793  
QY 510 ArgArgGlySerIleValSerLeuPro-----GlyGlyGlyAspValPheGlu 525  
Db 794 CTCGACGGGGACGTAGTCAACCGTACCCGGCAAGGAGCAAGCGCGGTGAGTTG----- 847  
QY 526 GlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly 545  
Db 847 ----- 847  
QY 546 LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys 565  
Db 847 ----- 847  
QY 566 GlySerArgTrpLysAspLeuPhe-----GluProGlyValArgArgAlaLeuLeu 582  
Db 848 ---CAGGTGTGAAGAAAGCTCATCCTGTCCCCGACCCCGGTGTCCGACGCATCTGCTC 904  
QY 583 ValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyr 602  
Db 905 TCGGCCGTGGGTCTCCACTTCTTCCAGCAGCGTCTTGGCAGCGACTCCTCGTCTCCAGTAC 964  
QY 603 ThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSer 622  
Db 965 AGCGCCCGCCTGTTCAAGAGCGCGGGGATCACCCGACGACAACAAGTCCTGGCGTCACC 1024  
QY 623 SerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle 642  
Db 1025 TCGCGG-----GTGGGCGTGACCAAGACGTTCTTCATCCTG----- 1060  
QY 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIle 662  
Db 1061 ---GTGGCCACGTTCTCTGCTGGACCGCGCGGGCGTCTGGCTCTGCTGATCAGCACG 1117  
QY 663 ProIleLeuIleAlaSerIleValIleLeu-----ValValSerAsnLeu 677  
Db 1118 GCGCGGATGATTGTCTCGTCTCATCTGCCTCGGGTCGGGGCTCACCCGTCGGGGCATCAC 1177  
QY 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697  
Db 1178 CCGGACACCAAGGTGCGGTGGGCGGTGCGCCCTGTGCATCGCGTCAACCTGTCTCATATC 1237  
QY 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro 717  
Db 1238 GCCTTCTTCTCCATCGGCCCTCGGGCCCATCACGGGCGGTGTACACCTCGGAAATATCCCCG 1297  
QY 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737  
Db 1298 CTGACGTGCGCGCTGGGCTTCGCGGTGGGTGTGGCGAGCAACCGCGTCACCGAGCGCC 1357  
QY 738 IleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757  
Db 1358 GTCATCTCCATGACCTTCTCTGTCCTCTCCAAAGGCCATCACCATCGGGCGGACGCTTCTTC 1417  
QY 758 IleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThr 777  
Db 1418 CTCTACTCCGGCATCGCGCGGTTCGTTGGGTTTTTCTTCTTTCACGTGCCTCCCGGAGACA 1477  
QY 778 LysGlyMetProLeuGluValIleThrGluPhePhe-----Ala 790  
Db 1478 CGCGGCCGACGCTGGAGGAGATGGGCAAGCTGTTCCGCATGCCAGACACGGGCAATGGCT 1537





Db	920	-----TGAAGGAGCTCATCTTTTCGCCGACCCAGCCATG	955
Qy	578	ArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsn	597
Db	956	CGGCGCATACTGTCGGCGCTCGGCATCCATTTCTTCAGCAGCGACGGGCTCCGAC	1015
Qy	598	GlyValLeuTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSer	617
Db	1016	TCCGTGCTCTATAGCCACGCGTGTCCAGAGCGCGGCATCACCGCGCAACACCAC	1075
Qy	618	LysPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeu	637
Db	1076	CTGCTCGGC-----GCCACATGCGCCATGGGGTTCATGAAGACGCTCTTC	1120
Qy	638	MetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeu	657
Db	1121	ATCCTG-----GTGGCCACGTTCCAGCTCGACCGCGTCGGCAGGCGCGCTG	1168
Qy	658	LeuLeuGlyThrIleProIleLeuIleAlaSerLeuVal-----IleLeu	672
Db	1169	CTGCTGACCAGCAGCGCCGCGCATGTCGCCTGTCTCATCGGCCTCGGACGGGCTCACC	1228
Qy	673	ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer	692
Db	1229	GTCGTGGGTGGCACCAGCGACGCCAAGGTCCCGTGGGCCATCGGCCTGTGCATCGTGTCC	1288
Qy	693	ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys	712
Db	1289	ATCTTGGCCTACGTCTCTTCTCCATCGGCCTCGGGCCCTCACCCAGCGTGTACACC	1348
Qy	713	AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe	732
Db	1349	TCGGAGGTCTTCCACACTGCGGGTGGCGCGCTGGGCTTCGCGTGGGCACGTGATGCAAC	1408
Qy	733	TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu	752
Db	1409	CGCGTCACCAAGCGCGGTCTCCATGTCTCTTCCTGTCTGTCCAAAGGCCATCACCATC	1468
Qy	753	AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu	772
Db	1469	GGCGGCAGCTTCTCTGTACCGCGGCATCGCGCGATAGGATGATTTTCTCTTCTCACC	1528
Qy	773	LysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePhe-----	789
Db	1529	TTCATTCCGGAGACGCGTGGCCTGCCGTCTCGAGGAGATAGGAAGCTTTTTCGGCATGACG	1588
Qy	790	-----AlaValGlyAlaLysGlnAlaAlaLys	799
Db	1589	GACACGGCGGTGGAAGCCCAAGACACCGCCACGAA	1624

Qy	244	ProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValIleuGlnLysLeu	263
		    :    :	
Db	686	CCCCGGTGGCTCGTCATGAAGGCCGTCTCGCGACGCCAAGTTGTGCTTGCCAAGACG	745
		    :    :	
Qy	264	ArgGlyLysAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGly	283
Db	745	-----	745
Qy	284	GlyAspThrSerIleGluGluTyriIleIleGlyProAlaThrGluAlaAlaAspAspLeu	303
Db	746	TCCGACACG-----	754
Qy	304	ValThrAspGlyAspLysGluGlnIleThrLeuTyrgLyProGluGluGlyGlnSerTrp	323
		    :	
Db	755	-----CCGGAAGAGCGCCGCAGCGC	775
Qy	324	IleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArg	343
Db	776	ATCGCC-----	781
Qy	344	HisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGly	363

Db 781 ----- 781  
Qy 364 SerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsn 383  
Db 781 ----- 781  
Qy 384 PheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGlu 403  
Db 781 ----- 781  
Qy 404 AsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAsp 423  
Db 782 -----GACATTAAAGACTGCCGC----- 799  
Qy 424 AsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleVal 443  
Db 799 ----- 799  
Qy 444 HisHisGlyHisArgGlySerAlaLeuSerMetArgGlnThrLeuLeuGlyGluGly 463  
Db 799 ----- 799  
Qy 464 GlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSer 483  
Db 799 ----- 799  
Qy 484 GluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGln 503  
Db 799 ----- 799  
Qy 504 GluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspVal 523  
Db 800 ---GGCATCCCTCTGGGCCTCGACGGCGAGCTGGTCCCCGTGCCCAA----- 844  
Qy 524 PheGluGlySerGluPheValHisAlaAlaAlaLeuValSerGlnSerAlaLeuPheSer 543  
Db 845 -----AAC 847  
Qy 544 LysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAla 563  
Db 848 AAAGGAAGCAGCAGGAGAGCGCGTTTGAAGGACCTCATCCTGTGACCGACCATAGCC 907  
Qy 564 AlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuVal 583  
Db 908 -----ATGGGCCACATCCTCATCGCG 928  
Qy 584 GlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThr 603  
Db 929 GGAATCGGCATCCACTTCTTCCAGCAGTCTTCGGGCATCGACGCCGTGCTCTACAGC 988  
Qy 604 ProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSer 623  
Db 989 CCGCTAGTTTCAAGAGCGCGGCATCAGC-----GGCGACAGCCGT 1030  
Qy 624 ---AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIle 642  
Db 1031 CTCCGCGGCACCAACCGTGGCGGTTCGGGGCCACCAATACGGTCTTTCATCCTG----- 1081  
Qy 643 GlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIle 662  
Db 1082 ---GTGGCCACCTTCTCTCGACCCGATCCGCCGGCGCGCTGGTGCTGACCAACG 1138  
Qy 663 ProIleLeuIleAlaSerLeuVal-----IleLeuValValSerAsnLeu 677  
Db 1139 GCGCGCATGCTCGTCTCCTTAGTGGCCTCGCGACGGGCTCACCGTCTCATCGCCGCCAC 1198  
Qy 678 IleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPhe 697  
Db 1199 CCGGACGAGAGATCACCTGGGCCATCGTCTGTGCATCTTCTGTCATCATGGCCTACGTG 1258  
Qy 698 CysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhePro 717  
Db 1259 GCCTTCTTCTCCATCGGCCTCGGGCCCCCATCACGTGGGTGTACAGCTCGGAGATCTTCCCG 1318

Qy 718 ThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIle 737  
Db 1319 CTGCACGTGCGCGCGTGGGTGCTCCCTGGCGGTGGCGCTCAACCGCCTGACCAAGCGG 1378  
Qy 738 IleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSer 757  
Db 1379 GTGATCTCCATGACCTTCATTTCGCTGTCCAAGGCCATGACCATCGCGCGCCTTCTTC 1438  
Qy 758 IleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThr 777  
Db 1439 CTCTTCGCGGCATCGCCTCATTCGCATGGGTGTCTTCTTCGCCTACCTGCCGAGACC 1498  
Qy 778 LysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla-----LysGln 795  
Db 1499 CGCGCGCGCACGCTGGAGGACATGAGCTCGCTGTTCGGCAACACGCGCCACCAAGCAG 1558  
Qy 796 AlaAlaAlaLysAla 800  
Db 1559 GGCGCGCGGAAGCC 1573

RESULT 12

US-09-489-039A-4762  
; Sequence 4762, Application US/09489039A  
; Patent No. 6610836

GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 4762  
; LENGTH: 1431  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-4762

Alignment Scores:

Pred. No.: 5.62e-37 Length: 1431  
Score: 447.00 Matches: 158  
Percent Similarity: 33.68% Conservative: 98  
Best Local Similarity: 20.79% Mismatches: 186  
Query Match: 10.97% Indels: 318  
DB: 4 Gaps: 18

US-10-051-909-32 (1-800) x US-09-489-039A-4762 (1-1431)

Qy 42 SerGlnGluProValThrSer---AspAspIleLeuGluAsp---LysMetSerGly--- 58  
Db 4 TCGACAGAATCAATAACTCAACTGGAGGGCGTTATGCCCTGACAAACAAAAACAAGGGCGT 63  
Qy 59 -----AlaValLeuValAlaIleValAlaSerIleGlyAsnLeuGln 73  
Db 64 TCGAACAAAGACTATGACGTTCTTCGTCTGTTCCTCGCGCGCTGGCTGCTGTTTC 123  
Qy 74 GlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeu 93  
Db 124 GGCCTTGATATCGGTGTTATTGCGGGTGCCTTACCCCTTATTGCCAATGAGTTCAGATT 183  
Qy 94 GlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIle 113  
Db 184 TCCGCCAC-----ACCCAGGAGTGGTGGTCAGCTCCATGATGTTTCGGGGCTGCCGTC 237  
Qy 114 ValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgPrometLeuIleLeu 133  
Db 238 GGCGCGGTCCGACGGCTGGCTCTCTTTCAAACCTGGGCCGAAAAAGAGCCTGATGATC 297  
Qy 134 SerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrVal 153

Db 298 GGGCCATCCTCTTCGTCGCCGTTTCGCTGTCTCTCTCGCCGCGCCGCAAAACGTCGAGATC 357  
QY 154 LeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValPro 173  
Db 358 CTGCTGGTTTCCCGTGTGCTGCGCTCGGCTGGCGTGGCGTGCCTCATATACGGCTCCG 417  
QY 174 LeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGln 193  
Db 418 CTGTATCTGTGGAAATCGCCAGAAAAATTCGGCGCAGTATGATTTCATGTATCCAG 477  
QY 194 PheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213  
Db 478 CTGATGATCACCATCGGGATCCTTGGCGCTAT-----CTCTCTGACACCGCTTTCAGC 531  
QY 214 ProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233  
Db 532 TACAGCGGCGCATGGCGTGATGCTCGGGTTATCATATCCGGCGGTTTGTGCTG 591  
QY 234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253  
Db 592 --ATCGCGGTATCTTCTCTCGCGACAGCCCGCTGGTTTCGCCGCAAAACGTCGCTTT 648  
QY 254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273  
Db 649 GTCGATCGGAACGCTGCTGCGCTGCGC-----GATACCAGCGCCGAA----- 681  
QY 274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIle 293  
Db 682 -----GATACCAGCGCCGAA----- 696  
QY 294 GlyProAlaThrGluAlaAlaAspLeuValThrAspGlyAspLysGluGlnIleThr 313  
Db 696 ----- 696  
QY 314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333  
Db 696 ----- 696  
QY 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353  
Db 696 ----- 696  
QY 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373  
Db 696 ----- 696  
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
Db 696 ----- 696  
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413  
Db 697 ---GCGAAACGCGAGCTC----- 711  
QY 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433  
Db 712 -----GATGAAATCCGTGAAAGCCTGAAGGTAAACAG 744  
QY 434 AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeuSer 453  
Db 745 TCC----- 747  
QY 454 MetArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerThrAspIleGly 473  
Db 747 ----- 747  
QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493  
Db 748 ---GGCTGGTCGCTG----- 759  
QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513  
Db 760 -----TTTAAA----- 765

QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAala 533  
Db 766 -----GACAAACAGCAACTTC----- 780  
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553  
Db 780 ----- 780  
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573  
Db 780 ----- 780  
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPhe 593  
Db 781 -----CGCGCGCGGTGTTCTCGGCATCCTGCTGCAGGTGATGCAACAGTTC 828  
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613  
Db 829 ACCGGGATGAACGTCATCATGTACTACGCGCCGAAGATCTTTGAGCTGGCGGTTATGCC 888  
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633  
Db 889 AACACCACCTGAGCAATGTGG-----GGGACAGTGATCGTCGGTCTGACT 933  
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653  
Db 934 AACGTGCTGGCCACCTTTATCGCCATCGGT-----CTGGTCGACCGCTGGGGC 981  
QY 654 ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673  
Db 982 CGTAAA-----CCGACGCTGATCCTTGGCTTTATCGTATGGCC 1020  
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHis----- 685  
Db 1021 GCGGGAATGGCGGTC--CTGGGTACCATGATGCACATCGGCATTCACTCCTCTACCGCC 1077  
QY 686 AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly 705  
Db 1078 CAGTACATCGCCGCTCTGATGCTGTGATGTTTCGCGGTTTCGCCATGAGCGCCGCGC 1137  
QY 706 ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle 725  
Db 1138 CCACTGATTGGGTACTGTGCTCCGAAATCCAGCGCTGAAAGCGCGGACTTCGGTATC 1197  
QY 726 AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal 745  
Db 1198 ACCTGCTCCACAGCGACCAACTGGATTGCCAACATGATTGTGGCGCCACCTTCTTGACC 1257  
QY 746 MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValCysLeuIle 765  
Db 1258 ATGCTCAACTCGCTGGCGACGCGCAATACCTTCTGGGTGTACGGCGGTCTGAACGTGCTG 1317  
QY 766 SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle 785  
Db 1318 TTTATCCTGCTGACGCTGTGGCTGATCCCGAAACCAAAACAGTCTCGCTGGAACATATT 1377  
RESULT 13  
US-09-489-039A-4731  
; Sequence 4731, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 4731  
; LENGTH: 1545



; TYPE: DNA		
; ORGANISM: Klebsiella pneumoniae		
US-09-489-039A-4731		
Alignment Scores:		
Pred. No.:	3.15e-36	Length: 1545
Score:	440.50	Matches: 164
Percent Similarity:	34.05%	Conservative: 105
Beat Local Similarity:	20.76%	Mismatches: 206
Query Match:	10.81%	Indels: 315
DB:	4	Gaps: 16
US-10-051-909-32 (1-800) x US-09-489-039A-4731 (1-1545)		
Qy	25	ProSerValValLeuAlaLeuProGlyProLeuProProAlaSer----- 39
Db	37	CCATATCTTTGGCTTATCCTGCCTGACCCATTCCACCAGTCAAGTCACTTTCGTCTTA 96
Qy	40	-----CysSerSerGlnGlu-----ProValThrSerSerAspIleLeuGlu 53
Db	97	TGTCCTTACTGTGTGGCAGGAAAAAATGACTTCAATCAGTAACGACTCTACATTATCG 156
Qy	54	AspLysMetSerGlyAlaVal-----LeuValAlaIleValAlaSerIle 68
Db	157	CCGCGGACGCAACGTGATACCCGGCGGATGAAGTGGTTTGTTCCTATCGCGCGGTG 216
Qy	69	GlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLys 88
Db	217	GCTGGCTTGCTCTTTGGCCTGGATATCGCGGTGATATCCGGAGCGTTCCTTTATAACC 276
Qy	89	LysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeu 108
Db	277	GACCATTTACACTTATCCAGCCAG-----CTTCAGGAGTGGTGTGTAGCAGCATGATG 330
Qy	109	IleGlyAlaThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArg 128
Db	331	TTGGGGCGCGGATAGGCGGCTGTTTAAACGGCTGGCTGTCTTTCCGCTTGCCGATAA 390
Qy	129	ProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSer 148
Db	391	TACAGCCTGATGGGGGGCGGTACTCTTTGTGGCGGCTCTATCGGATCCGCTTTTGCC 450
Qy	149	ProAsnValTyrValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAla 168
Db	451	GCCAGCTGGAGGTGCTGTGGTGGCCGCGTGTGTGGCGTGGCAGTCGGGATTGCC 510
Qy	169	ValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeu 188
Db	511	TCTTATACCGCGCGCTGTACCTCTCCGAGATGGCCAGCGAGACGTGCGGGAATAAG 570
Qy	189	AsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPhe 208
Db	571	ATCAGTATGTATCAGCTGATGGTCCCTTGGCATTTGTCTGGCGTTT-----CTTTCC 624
Qy	209	GlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIlePro 228
Db	625	GATACCGCCTTTAGCTACAGCGGTAACTGGCGGCCATGCTGGGCGTGTGGCGTCCG 684
Qy	229	SerLeuPhePhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuVal 248
Db	685	GCGGTGATCCTGATCATTTCTGGTCGTCTTT---TTGCCGAACAGCCGCGCTGGCTGGCG 741
Qy	249	SerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAsp 268
Db	742	GAGAAGGAGCGCCATATCGAAGCGGAAGTGTGCGGATGCTGCGC----- 789
Qy	269	ValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIle 288
Db	790	-----GATACCTCG--- 798
Qy	289	GluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAsp 308
Db	799	-----GAAAAGGCGCGGACGAGCTT----- 819
Qy	309	LysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSer 328
Db	819	----- 819
Qy	329	LysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetVal 348
Db	820	-----AACGAGATC 828
Qy	349	AsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsn 368
Db	829	CGTGAGAGCCTCAAGCTGAAGCAG----- 852
Qy	369	MetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPhe 388
Db	852	----- 852
Qy	389	SerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAsp 408
Db	852	----- 852
Qy	409	AspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerPro 428
Db	852	----- 852
Qy	429	LeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArg 448
Db	852	----- 852
Qy	449	GlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSer 468
Db	852	----- 852
Qy	469	SerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGlu 488
Db	853	-----GGCGGTTGG----- 861
Qy	489	AsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGly 508
Db	861	----- 861
Qy	509	SerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGlu 528
Db	861	----- 861
Qy	529	PheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGlu 548
Db	862	-----GCGTTGTTTAAG----- 873
Qy	549	ProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArg 568
Db	873	----- 873
Qy	569	TrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGln 588
Db	874	-----ATCAATCGTAACGTGCGCGCGGTGTCTCTGGGCATGCTGTGTCAG 921
Qy	589	IleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGlu 608
Db	922	GCGATGCAGCAATTCCCGGCATGAACATCATCATGTACTATATGCGCGGTATCTTTAAA 981
Qy	609	GlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIle 628
Db	982	ATGCGC-----GGCTTTACCACTACTGAACAGCAGATG 1014
Qy	629	LeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeu 648
Db	1015	ATCGCCACCCTGGTGGTGGCGCTGACCTTTATGTTTGCACCTTTATTGCGGTGTTACA 1074
Qy	649	MetAspLeuSerGlyArgArg---PheLeuLeuLeuGlyThrIleProIleLeuIleAla 667
Db	1075	GTGGATAAAGCGGGCCGCAAGCCAGCGCTGAAATCGGCTTTAGCGTGTATGGCGTGGGC 1134

```
QY      668 SerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeu 687
Db      1135 ACTCTGGTGTGGGCTACTGCCTGATGCAGTTCCACAATGGCACC---GCATCCAGCGGC 1191
QY      688 LeuSerThrValSerValIleValTyrPheCysCys-----PheValMetGlyPhe 704
Db      1192 CTCTCTGGCTCTCCGTCGGCATGACCATGATGTTGTTATGCCGGTATGCCGATGAGCGCG 1251
QY      705 GlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCys 724
Db      1252 GCGCCGGTGGTGTGGATCCTCTGCTCCGAGATCCAGCGCTAAATGCCGCGACTTCGGT 1311
QY      725 IleAlaIleCysAlaPheThrPheThrPheIleGlyAspIleIleValThrTyrSerLeuPro 744
Db      1312 ATCACCTGCTCGACCAACCACCACTGGGTGTGCAACATGATCATCGCGCGCACTTTCCTG 1371
QY      745 ValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeu 764
Db      1372 ACGCTGTCTGACGCGATTTGGCGCGCGCGCACCTTCTGGCTCTACACGGCGCTCAACGTG 1431
QY      765 IleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluVal 784
Db      1432 GCCTTTATCGGCATCACCTTCTGGCTGATCCCGGAACCAAGAATGTCAACCTCGAGCAC 1491
QY      785 IleThrGluPhePheAlaValGlyAlaLys 794
Db      1492 ATTGAGCGCACTGATGGCGGCGGAGAAG 1521

RESULT 14
US-09-489-039A-4560
; Sequence 4560, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4560
; LENGTH: 1506
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4560

Alignment Scores:
Pred. No.:      8.15e-35      Length:      1506
Score:          427.00      Matches:      155
Percent Similarity: 33.29%      Conservative: 104
Best Local Similarity: 19.92%      Mismatches: 201
Query Match:      10.48%      Indels:      318
DB:              4          Gaps:      17

US-10-051-909-32 (1-800) x US-09-489-039A-4560 (1-1506)
QY      33 GlyProLeuProProAlaSerCysSerSerGlnGluPro-----ValThrSer 48
Db      19 GGCCCG---CCTGAACCCCTACTGTTCCCTGCGTTTCCCTTATAGAGGAATCATTATGAAC 75
QY      49 AspAspIleLeuGluAspLysMetSerGlyAlaValLeuValAlaIleValAlaSerIle 68
Db      76 AACGCGCAGACACATCTGAAATGGGTACGTCTGGACGATTTGTCTGTGTCGCCCGCTGC 135
QY      69 GlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAlaAlaValLeuTyrIleLys 88
Db      136 GGTGGTTTACTGTTTGGCTATGACTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 195
QY      89 LysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeu 108
Db      196 GCCTGGTTTCAATT---ACGACCCCGCGCAGTCCGCGTGGCGGATGATGATGATGATGATGATG 252
```

```
QY      109 IleGlyAlaThrIleValThrThrPheSerGlyProLeuSerAspSerIleGlyArgArg 128
Db      253 TTGGGCTGTATTTTCGGCGCATTAATTCGGGATGGTGGCGACACAACATGGGGCGCAAG 312
QY      129 ProMetLeuIleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSer 148
Db      313 CTGCCATTAAATCTTCCGCCGTGCTGTTCAGCGCGTCCGCCCTGGGGGACGGCGGTCCGCC 372
QY      149 ProAsnValTyrValLeuLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAla 168
Db      373 AGTCATTTCGATATGTTTGTGGTTTACCGCATTTGTGGCGCGGTAGGATTTGGTCTGGCT 432
QY      169 ValThrLeuValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeu 188
Db      433 TCCGCTCTCAGCCCGCTTTACATTGCCGAATCAGCCCGGAGAGAAAAGAGGACGTTTTT 492
QY      189 AsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSerTyrCysMetValPhe 208
Db      493 GTCGCGGTCAATCAGCTCACCATCGTGTATTGGCGTGTGGCCGCTCAGTTAATCAATCTG 552
QY      209 GlyMetSerLeuSerProSerPro----- 216
Db      553 ATGATTGCTGAACCGGTGGAGCCGGGGCGGACGACAGATGATTGTGGACAGCTGGAAT 612
QY      217 -----AspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db      613 GGCAGATGGGCTGGCGCTGGATGTTCCGTGCGGAATGTTGCCGCACTGGCGTCTCTG 672
QY      234 GlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet 253
Db      673 GTCCTG---ATGTTTTTGTCCCGAGTCCCGCGCTGGCTCATGAAGCCCGTAAACCG 729
QY      254 AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
Db      730 GAGCGCGCCGGGCTGCGCTGGAACG----- 756
QY      274 SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIle 293
Db      757 ----- 759
QY      294 GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db      760 GGTTCCTGCC----- 768
QY      314 LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet 333
Db      768 ----- 768
QY      334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353
Db      768 ----- 768
QY      354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373
Db      768 ----- 768
QY      374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393
Db      768 ----- 768
QY      394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413
Db      768 ----- 768
QY      414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433
Db      769 -----GACTATGCCGACAGGATCCTG----- 789
QY      434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453
Db      790 -----CGTGAAATCGCGCAT----- 804
```

```
QY 454 MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly 473
Db 805 -----ACCTG-----
QY 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493
Db 811 -----GAAAGGATAACAATAA-----
QY 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513
Db 828 -----
QY 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533
Db 828 -----
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553
Db 829 -----GTCCTCCTACGGCGCGCTGTGGCT-----
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573
Db 852 -----
QY 574 GluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593
Db 853 ---CCCAAGGTGAACCGATTGTGATCATTTGGCATGGTGCTCGCCATATTCCAGCAGTGG 909
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613
Db 910 TGTGGGATTAAACGTCACTTTAACTACGCGCAGGAGATTTTGGCCTCGCGGGG----- 963
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633
Db 964 -----TTCGATATTAAACAGCACGCTGAAATCGATCGTCCGACGGCGGTC 1008
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653
Db 1009 GTTAATCTGGTCTTTACCATTTGCG-----GCGCTGCCGTGGTGGATAAAATCGGT 1059
QY 654 ArgArg---PheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeu 672
Db 1060 CGCCGTAAATTAATAGTCTGTTGGCGCTTCGGGATTG-----ACGCTGATCTATGTG 1110
QY 673 ValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSer 692
Db 1111 CTGATCGCGCGCGCTTACGCCATGGGCATTATGGGTGGCGCGGTACTG-----TTGCTG 1164
QY 693 ValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCys 712
Db 1165 GTGCTGGCGCGGATTGCTATTTATGCGTGTGACCTGGCGCGGTGACCTGGGTGCTGCTG 1224
QY 713 AlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPhe 732
Db 1225 GCGGAGATTTTCCCAACCGCGTACGTGGACTCGCGATGTCTTTAGGTACCTCGCGCTG 1284
QY 733 TrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeu 752
Db 1285 TGGATCGCCTGTTTCCCTGTTAAACCTATACTTTCCCGTGTCTTAATGCCGGTCTGGCGCG 1344
QY 753 AlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeu 772
Db 1345 GCAGGAAGCTTCCTGCTGTATGGTGTATTATTCGGCTGCCGGCTACCTCTATATCCTGCGC 1404
QY 773 LysValProGluThrLysGlyMetProLeuGluValIleIleThrGluPhePheAla 790
Db 1405 AACGTCCCGGAAACGAAAGGTATCACCCTCGAAGCGCTGGAAGAGCAGCTGGCG 1458
```

RESULT 15

```
US-09-489-039A-2378
; Sequence 2378, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
```

```
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2378
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2378
```

Alignment Scores:

```
Pred. No.: 1.11e-34 Length: 1566
Score: 426.00 Matches: 163
Percent Similarity: 34.24% Conservative: 88
Best Local Similarity: 22.24% Mismatches: 176
Query Match: 10.45% Indels: 306
DB: 4 Gaps: 20
```

US-10-051-909-32 (1-800) x US-09-489-039A-2378 (1-1566)

```
QY 62 ValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAla 81
Db 172 ATCACTTAGTGGCCACCTTCGGCGGCTGCTTTTCGGTTACGACACCGGCGTTATCAAC 231
QY 82 AlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGluGly 101
Db 232 GCGGCTTTTCTCACTGAAGCAGTACATGGCGTG-----ACCCCAACCACCGAAGG 285
QY 102 LeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeu 121
Db 286 CTGTAATGAGCGTTCTGCTGGTGGCGCGCGCTGGGACGCGTGTGTGGCGGCAAAATT 345
QY 122 SerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePheSer 141
Db 346 GCCGATTACTTCGGCCGGCGTAAATATTACTTTCTCTCTCTTTGTGTTTCTCATCGGC 405
QY 142 GlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheValAsp 161
Db 406 GCGTTGCTTTCGGCGCGCGCCCGATATCACTACCTGCTGATGTCGCCGCTGCTG 465
QY 162 GlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaPro 181
Db 466 GGCTATGCCGTCGGCGGCGCTTCGGTCACCGCCCCACCTTTATTATTCAGAAAGTGGCGCT 525
QY 182 SerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPhe 201
Db 526 ACCGAGATGCGCGGCAAGCTGACCGGCTGAATGAAGTCGCCATCGTCATTGGCCAGCTG 585
QY 202 LeuSerTyrCysMet-----ValPheGlyMetSerLeuSerProSerProAsp---Trp 218
Db 586 GCGGCTTTGCCATCAACGCGCATCATTTGGCATCATCTGGGGCCATCTGCCGACGTCGTG 645
QY 219 ArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePhe 238
Db 646 CGCTATATGCTGCTGGTACAAGCGATTCCGGCCATCTCGCCATCTGCCTGTTTC---GTTGGCATGTGG 702
QY 239 TyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLys 258
Db 703 CGCGCGCGGAAAGCCCGCGCTGGCTGATCAGCAAAAATCGTCATGATGAAGCGCTGCAT 762
QY 259 ValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuGlu 278
Db 763 ATCCTCAAAACAAATCCGC-----
QY 279 GlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGlu 298
Db 781 -----CCGGCGGAACGC 792
```





GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 18:06:15 ; Search time 1016.41 Seconds  
(without alignments)  
4775.020 Million cell updates/sec

Title: US-10-051-909-32  
Perfect score: 4075  
Sequence: 1 IRGSLAVQTPTPDLRR.....PLEVITEFFAVGAKQAAKA 800

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074039\_14141/app\_query.fasta\_1.1678  
-DB=Published\_Applications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10051909 @CGN 1 1 953 @runat\_13042005\_074039\_14141  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*

4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_NEW\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*

14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*

17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*

18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:\*

19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*

20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*

21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*

22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4075	100.0	2777	13	US-10-051-909-31	Sequence 31, Appl
2	4059	99.6	2908	17	US-10-310-154-350	Sequence 350, Appl
3	3517	86.3	2824	13	US-10-051-902-1	Sequence 1, Appli
4	3517	86.3	2824	13	US-10-051-909-1	Sequence 1, Appli
5	2869	70.4	2031	17	US-10-425-114-32961	Sequence 32961, A
c	2835.5	69.6	2769	18	US-10-437-963-93472	Sequence 93472, A
6	2788.5	68.4	4330	18	US-10-425-115-70742	Sequence 70742, A
7	2674	65.6	2601	13	US-10-051-902-7	Sequence 7, Appli
8	2674	65.6	2601	13	US-10-051-909-7	Sequence 7, Appli
9	2584	63.4	2190	9	US-09-938-842A-1315	Sequence 1315, Ap
10	2584	63.4	2190	11	US-09-938-842A-1315	Sequence 1315, Ap
11	2571	63.1	2671	17	US-10-424-599-10324	Sequence 10324, A
12	2543	62.4	2405	17	US-10-425-114-7958	Sequence 7958, A
13	2483.5	60.9	2573	17	US-10-424-599-108435	Sequence 108435, A
14	2442.5	59.9	2271	17	US-10-425-114-29146	Sequence 29146, A
15	2275	55.8	2205	9	US-09-938-842A-2254	Sequence 2254, Ap
16	2275	55.8	2205	11	US-09-938-842A-2254	Sequence 2254, Ap
17	2245.5	55.1	2432	18	US-10-437-963-47424	Sequence 47424, A
c	1963	48.2	1959	18	US-10-437-963-92959	Sequence 92959, A
18	1763	43.3	2231	18	US-10-437-963-70909	Sequence 70909, A
19	1670.5	41.0	1692	13	US-10-051-902-9	Sequence 9, Appli
20	1670.5	41.0	1692	13	US-10-051-909-9	Sequence 9, Appli
21	1545	37.9	3205	17	US-10-424-599-77967	Sequence 77967, A
22	1424	34.9	1487	13	US-10-051-902-13	Sequence 13, Appl
23	1424	34.9	1487	13	US-10-051-909-13	Sequence 13, Appl
24	1392	34.2	1313	18	US-10-767-701-12595	Sequence 12595, A
25	1392	34.2	1313	18	US-10-767-701-12595	Sequence 12595, A
26	1252.5	30.7	1289	17	US-10-424-599-42000	Sequence 42000, A
27	1035	25.4	1412	17	US-10-260-238-747	Sequence 747, App
28	959	23.5	1200	18	US-10-425-115-159851	Sequence 159851, A
29	949.5	23.3	1259	17	US-10-424-599-101136	Sequence 101136, A
c	943	23.1	1009	13	US-10-051-902-15	Sequence 15, Appl
30	943	23.1	1009	13	US-10-051-909-15	Sequence 15, Appl
31	813	20.0	761	17	US-10-425-114-35910	Sequence 35910, A
32	793	19.5	650	17	US-10-425-114-3864	Sequence 3864, Ap
33	745.5	18.3	832	18	US-10-425-115-135446	Sequence 135446, A
34	663	16.3	435	18	US-10-767-701-18822	Sequence 18822, A
35	620	15.2	711	18	US-10-767-701-12614	Sequence 12614, A
36	620	15.2	751	17	US-10-260-238-2683	Sequence 2683, Ap
37	615.5	15.1	2253	18	US-10-437-963-49329	Sequence 49329, A
38	604	14.8	674	18	US-10-437-963-60088	Sequence 60088, A
39	600	14.7	778	17	US-10-260-238-2684	Sequence 2684, Ap
40	587.5	14.4	1374	17	US-10-369-493-47011	Sequence 47011, A
41	587.5	14.4	1518	19	US-10-332-815A-10	Sequence 10, Appl
42	580	14.2	583	18	US-10-021-323-11734	Sequence 11734, A
43	562	13.8	870	13	US-10-051-902-5	Sequence 5, Appli
44						
45						

ALIGNMENTS

RESULT 1  
US-10-051-909-31  
; Sequence 31, Application US/10051909  
; Publication No. US20020199217A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Helentjaris, Tim  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB1163 US CIP  
; CURRENT APPLICATION NUMBER: US/10/051,909  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 60/083,044  
; PRIOR FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 31  
; LENGTH: 2777

; TYPE: DNA					
; ORGANISM: Zea mays					
US-10-051-909-31					
Alignment Scores:					
Pred. No.:	0	Length:	2777		
Score:	4075.00	Matches:	800		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	13	Gaps:	0		
US-10-051-909-32 (1-800) x US-10-051-909-31 (1-2777)					
QY	1	IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArg	20		
Db	12	ATTCGAGCGGCTCTTGCTTGAGTCCAAACGCCCTTCACCCCTGATCTGGACCGGAGG	71		
QY	21	GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProAlaSerCys	40		
Db	72	GAGCGGCTCCTCCGTACGTTGTCTTGCTTTGCCCTGGGCCCTCTTCCGCCCTGCTTCGTGT	131		
QY	41	SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal	60		
Db	132	TCTTCACAGGAGCCGGTGACCTCGACGATATCTTGGAGGACAAGATGTCGGGGCTGTT	191		
QY	61	LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle	80		
Db	192	CTTGTGCCCATAGTCGCCCTCCATCGGCAATCTATTGCAGGGGTGGACAAATGCCACCATC	251		
QY	81	AlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsnGluProThrValGlu	100		
Db	252	GCAGCTGCTGTTCTGTATATAAAGAAGGAATTTCAATTGCAAAATGAGCCCACTGTGGAG	311		
QY	101	GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrThrPheSerGlyPro	120		
Db	312	GGACTAATTGTGTCATGTCACTTATCGGCGCCACCATCGTTACTACATTTCTCGGGCCA	371		
QY	121	LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe	140		
Db	372	TTATCAGACTCGAATTGGCCGACGCCCTATGCTTATTCTCTCTTCAATTTCTGTACTTCTTC	431		
QY	141	SerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaArgPheVal	160		
Db	432	AGCGGCCTCATGTCTATGGTCTCCTAAATGCTATGTCTGTCTGTGTGGCACGCTTCGTA	491		
QY	161	AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla	180		
Db	492	GATGGATTGGTATTGGCTTGGCTGTACGCTTGTGCCCTTGTACATTTTCAGAAATAGCC	551		
QY	181	ProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyMet	200		
Db	552	CCTTCGGAGATTAGAGGTTTGCTGAATACACTACCACAAATTCAGTGGATCAGGAGGAATG	611		
QY	201	PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle	220		
Db	612	TTCTTGTCATACTGTCATGGTGTTTGGGATGTCCCTGTGCGCCATCACCCGATTGGAGAATT	671		
QY	221	MetLeuGlyValLeuAlaIleProSerLeuPhePhePheGlyLeuThrIlePheTyrLeu	240		
Db	672	ATGCTTGGTGTGCTCGCATACCTTCATTGTTCTTTTGGTTTGACAAATATTTTATCTT	731		
QY	241	ProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu	260		
Db	732	CCTGAATCTCCAAGATGGCTCGTTAGCAAAGGTGGATGGCAGAGGCCAAAAAGGTGTG	791		
QY	261	GlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuLeuGluGlyLeu	280		
Db	792	CAAAAGTTACGGGGGAAAGACGATGTCTCAGGTGAATTGTCCCTTCTTCTCGAAGGGTGTG	851		
QY	281	GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla	300		
Db	852	GAGGTTGGAGGAGACACTTCCATTGAAGAGTACATCATTTGGACCTGCCACCGAGGCAGCC	911		

QY	301	AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly	320
Db	912	GATGATCTTTGTTACTGACGGTGATAAGGAACAAATCACACTTTATGGGCTGAAGAAGGC	971
QY	321	GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu	340
Db	972	CAGTCATGGATTGCTCGACCTTCTAAGGGACCCCATCATGCTTGGAAAGTGTGCTTCTTT	1031
QY	341	AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr	360
Db	1032	GCATCTCGTCATGGGAGCATGGTGAACCCAGAGTGTACCCCTTATGGATCCGATTGTGACA	1091
QY	361	LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu	380
Db	1092	CTTTTGTGGTAGTCTCCATGAGAAATATGCCTCAAGCTGGAGGAAGTATGAGGAGCACATTG	1151
QY	381	PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp	400
Db	1152	TTTCCAAACTTTTGGAAAGTATGTTCAAGTGTACAGATCAGCATGCCAAAAATGAGCAGTGG	1211
QY	401	AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp	420
Db	1212	GATGAAGAGAATCTTCATAGGGATGACGAGGAGTACGCATCTGTATGGTGCAGGAGGTGAC	1271
QY	421	TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys	440
Db	1272	TATGAGGACAAATCTCCATAGCCCATTGCTGTCCAGGCAGGCAACAGGTGCGGAAGGGAAG	1331
QY	441	AspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu	460
Db	1332	GACATTGTGCACCATGGTCACCGTGGAAAGTCTTTGAGCATGAGAAGGCCAAACCCCTCTTA	1391
QY	461	GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrp	480
Db	1392	GGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGTGGGGATGGCAGCTTGCTTGG	1451
QY	481	LysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyr	500
Db	1452	AAATGGTCAGAGAAGGAAGGTGAGAAATGGTAGAAAGGAAGGTGTTTCAAAAGAGTCTAC	1511
QY	501	LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly	520
Db	1512	TTGCACCAAGAGGGAGTTCCTGGCTCAAGAAAGGGGCTCAATTGTTTTCACCTTCCCGTGGT	1571
QY	521	GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla	540
Db	1572	GGCGATGTTTTTGAGGGTAGTGAGTTTGTACATGCTGCTGCTTTTAGTAAGTCAGTCAGCA	1631
QY	541	LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer	560
Db	1632	CTTTTCTCAAAGGGTCTTGTGTGAACCAACGCATGTCAGATGCTGCCATGGTTCACCCATCT	1691
QY	561	GluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAla	580
Db	1692	GAGGTAGCTGCCAAAGGTTTCAAGTTGGAAAGATTGTTTGAACCTGGAGTGAGGCGTGCC	1751
QY	581	LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu	600
Db	1752	CTGTTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTTGTGGAATAAACCGTGTCTGTG	1811
QY	601	TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly	620
Db	1812	TACTATACCCACAAATTCCTTGAGCAAGCTGGTGTGGCAGTTATTCTTTCCAAATTTGGT	1871
QY	621	LeuSerSerAlaSerAlaSerIleLeuIleSerSerSerLeuThrThrLeuLeuMetLeuPro	640
Db	1872	CTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACTACCTTACTACTATGTTCCT	1931
QY	641	CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly	660
Db	1932	TGCATTGGCTTTGCCATGCTGCTTATGGATCTTTCCGGAAGAAGGTTTTTGTGCTAGGC	1991



```

Qy      661 ThrileProIleLeuIleAlaSerLeuValIleLeuValSerAenLeuIleAspLeu 680
      |||
Db      1992 ACAATTCCAATCTTGATAGCATCTCTAGTTATCCTGGTTGTGTCCCAATCTAATTGATTG 2051

Qy      681 GlyThrLeuAlaHisAlaLeuSerThrValSerValIleValTyrPheCysCysPhe 700
      |||
Db      2052 GGTACACTAGCCCATGCTTTGCTCTCCACCGTCAGTGTATTATCGTCTACTTCTGCTGCTTC 2111

Qy      701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720
      |||
Db      2112 GTTATGGGATTGTGTCCTCCATCCCCAACATTTTATGTGCAGAGATCTTTCCAACCAAGGTT 2171

Qy      721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThr 740
      |||
Db      2172 CGTGGCCTCTGTATTGCCATTTGTGCCCTTACATTTCTGGATCGGAGATATCATCGTCACC 2231

Qy      741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760
      |||
Db      2232 TACAGCCTTCCTGTGATGCTGAATGCTATTGGACTGGCGGGTGTTTTTCAGCATATATGCA 2291

Qy      761 ValValCysLeuIleSerPheValPheLeuLysValProGluThrLysGlyMet 780
      |||
Db      2292 GTCGTATGCTTGATTTCTTTGTGTCGTCCTTAAAGTCCCTGAGACAAAGGGGATG 2351

Qy      781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800
      |||
Db      2352 CCCCTTGAAGTTATTACCGAATTCCTTGCAAGTTGGTGCGAAGCAAGCGGCTGCAAAAGCC 2411

RESULT 2
US-10-310-154-350
; Sequence 350, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinzhao
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanquo
; APPLICANT: Xu, Nanfei

```

```

; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 350
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (331)..(2565)
; OTHER INFORMATION:
US-10-310-154-350

Alignment Scores:
Pred. No.:      0      Length:      2908
Score:          4059.00      Matches:      797
Percent Similarity: 99.75%      Conservative: 1
Best Local Similarity: 99.62%      Mismatches: 2
Query Match:      99.61%      Indels: 0
DB:              17      Gaps: 0

US-10-051-909-32 (1-800) x US-10-310-154-350 (1-2908)

Qy      1 IleArgSerGlySerTrpLeuAlaValGlnThrProPheThrProAspLeuAspArgArg 20
      |||
Db      166 ATTCGGAGCGGCTCTTGGCTTGCACTCCAGACGCCCTTCACCCCTGATCTGGACCGGAGG 225

Qy      21 GluArgLeuLeuProSerValValLeuAlaLeuProGlyProLeuProProAlaSerCys 40
      |||
Db      226 GAGCGGCTCCTCCGTCAGTTGTTCTTGCTTTGCCTGGGCCCTCTTCCGCCCTGCTTCGTGT 285

Qy      41 SerSerGlnGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaVal 60
      |||
Db      286 TCITTCACAGGAGCGGTGACCTCGACGATATCTTTGGAGGACAAGATGTGGGGGCTGTT 345

Qy      61 LeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIle 80
      |||
Db      346 CTTGTCGCCATAGTCGCCTCCATCGGCAATCTATTGCAGGGTGGACAATGCCACCATC 405

Qy      81 AlaAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnGluProThrValGlu 100
      |||
Db      406 GCAGCTGCTGTTCTGTATATAAAGAAGGAATTTCAATTGCAAAATGAGCCCACTGTGGAG 465

Qy      101 GlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyPro 120
      |||
Db      466 GGACTAATTGTGTCAATGTCACTTATCGGCGCCACCATCGTTACTACATTTCTCCGGGCCA 525

Qy      121 LeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyrPhePhe 140
      |||
Db      526 TTATCAGACTCGATTGGCCGACGCCCTATGCTTATTCTCTTCAATTCTGTACTTCTTC 585

Qy      141 SerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeuAlaArgPheVal 160
      |||
Db      586 AGCGGCCTCATCATGCTATGGTCTCTTAATGTCTATGTCTCTGTGTGGCAGCGCTTCGTA 645

Qy      161 AspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAla 180
      |||
Db      646 GATGGATTGGTATTGGCTTGGCTGTACGCTTGTGCCTTTGTACATTTCAGAAATAGCC 705

Qy      181 ProSerGluIleArgGlyLeuLeuAenThrLeuProGlnPheSerGlySerGlyGlyMet 200
      |||
Db      706 CCTTCGAGATTAGAGGTTTGGTGTGCTGTCACGCTTGTGCCTTTGTACATTTCAGAAATG 765

Qy      201 PheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIle 220

```

Db 766 |||||TCTGTGCATGCTGATGGTGT|||TGGGATGTCCTGTGCCATCACCGATTGGAGAAATT 825  
QY 221 MetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeu 240  
Db 826 ATGCTTGGTGTGTCGCGATACCTTCATTGTTCTCTTTGGTTTGACAAATATTTATCTT 885  
QY 241 ProGluSerProArgTyrLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeu 260  
Db 886 CCTGAATCTCCAAGATGGCTCGTTAGCAAAGGTCGATGGCAGAGGCCAAAAGGTGTG 945  
QY 261 GlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeuLeuGluGlyLeu 280  
Db 946 CAAAAGTTACGGGGGAAAGACGATGTCTCAGGTGAATTGTCCCTTCTTCTCGAAGGGTTG 1005  
QY 281 GluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAla 300  
Db 1006 GAGGTTGGAGGAGACACTTCCATTTGAAGAGTACATCATTTGGACCTGCCACCGAGCGAGCC 1065  
QY 301 AspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGly 320  
Db 1066 GATGATCTTGTACTGACGGTGTATAAGGAACAAATCACACTTTATGGGCCTGAAGAAGGC 1125  
QY 321 GlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeu 340  
Db 1126 CAGTCATGGATTGCTCGACCTTCCAAAGGGACCCAGCATGCTTGGAAAGTGTCTTCTCTT 1185  
QY 341 AlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThr 360  
Db 1186 GCATCTCGTCATGGGAGCATGGTGAACACAGAGTGTAACCCCTTATGGATCCGATGTGACA 1245  
QY 361 LeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeu 380  
Db 1246 CTTTTTGGTAGTGTCCATGAGAAATATGCCTCAAGCTGGAGGAAGTATGAGGAGCACATTG 1305  
QY 381 PheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrp 400  
Db 1306 TTTTCCAAACTTTGGAAGTATGTTCAAGTGTACAGATCAGCATGCCAAATGAGCAGTGG 1365  
QY 401 AspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAsp 420  
Db 1366 GATGAAGAGAAATCTTCATAGGGATGACGAGGAGTACCGCATCTGATGGTGCAGAGGTGAC 1425  
QY 421 TyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLys 440  
Db 1426 TATGAGGACAAATCTCCATAGCCCATGCTGTCCAGGCAGCAACAGGTGCGGAAGGGAAG 1485  
QY 441 AspIleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeu 460  
Db 1486 GACATTGTGCACCATGCTGTCACCGTGGAAAGTGCTTTGAGCATGAGAAGGCAAAAGCCTCTTA 1545  
QY 461 GlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrp 480  
Db 1546 GGGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGGTGGGGGATGGCAGCTTGTCTGG 1605  
QY 481 LysTrpSerGluLysGluGlyAsnGlyArgLysGluGlyGlyPheLysArgValTyr 500  
Db 1606 AAATGGTCAGAGAAGGAAGGTGAGAAATGGTAGAAAAGGAAGGTGGTTTCAAAAAGAGTCTAC 1665  
QY 501 LeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly 520  
Db 1666 TTGCACCAAGAGGAGTTCTTGCTCAAGAAAGGGCTCAATTGTTTCACTTCCCGGTGGT 1725  
QY 521 GlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAla 540  
Db 1726 GCGGATGTTCTTGAGGGTAGTGAGTTTGTACATGCTGTCTGTAGTAAGTCAGTCAGCA 1785  
QY 541 LeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSer 560  
Db 1786 CTTTTTCAAAGGGTCTTGCTGAACCAACCGATGTGATGCTGCAATGTTTCAACCCATCT 1845  
QY 561 GluValAlaAlaLysGlySerArgTyrLysAspLeuPheGluProGlyValArgArgAla 580  
|||||

Db 1846 GAGGTAGCTGCCAAAGGTTCCAGTTGGAAAGATTGTGTTTGAACCTGGAGTGAGCGTGCC 1905  
QY 581 LeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeu 600  
Db 1906 CTGTTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTTGTGGAATAAACGGTGTCTG 1965  
QY 601 TyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGly 620  
Db 1966 TACTATACCCCAAAATCTTTGAGCAAGCTGGTGTGGCAGTTATCTTTCCAAATTTGGT 2025  
QY 621 LeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrLeuLeuMetLeuPro 640  
Db 2026 CTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACTACCTTACTAATGCTTCT 2085  
QY 641 CysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGly 660  
Db 2086 TGCATTGGCTTTGCCATGCTGCTTATGGATCTTTCCGGAAGAGGTTTTTGTCTAGGC 2145  
QY 661 ThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeu 680  
Db 2146 ACAATTCCAATCTTGATAGCATCTCTAGTTATCCTGGTTGTGTCCAATCTAATTGATTG 2205  
QY 681 GlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPhe 700  
Db 2206 GGTACACTAGCCCATGCTTTGCTCTCCACCGTCAGTGTATCGICTACTTCTGCTGCTTC 2265  
QY 701 ValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgVal 720  
Db 2266 GTTATGGGATTTGGTCCCATCCCAACATTTTATGTGCAGAGATCTTCCAACCAAGGGTT 2325  
QY 721 ArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyrIleGlyAspIleIleValThr 740  
Db 2326 CGTGGCCTCTGTAATTGCCATTGTGCTTTCACCTTTCATCTGGATCGGAGATATCATCGTCACC 2385  
QY 741 TyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAla 760  
Db 2386 TACAGCCTTCTCTGATGCTGAATGCTATTGGACTGGCGGGTGTTCAGCATATATGCA 2445  
QY 761 ValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMet 780  
Db 2446 GTCGTATGCTTGATTTCTTGTGTTCGTCTTCTTAAGGTCCCTGAGACAAAAGGGGATG 2505  
QY 781 ProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAlaAlaLysAla 800  
Db 2506 CCCTTGAGGTTATTACCGAATTCTTTGCAGTTGGTGCGAAGCAAGCGGCTGCCAAAAGCC 2565

RESULT 3  
US-10-051-902-1  
; Sequence 1, Application US/10051902  
; Publication No. US20020178468A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tinney, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/10/051,902  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 1  
; LENGTH: 2824  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (29)  
; NAME/KEY: unsure  
; LOCATION: (622)  
; NAME/KEY: unsure

;	LOCATION:	(636)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
---	-----------	-------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

QY	274	SerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlyTyrIle	293
Db	898	GCTCTTCTAGTTGAAGGTTTGGGGTTCGGTAAAGATACACGTAATTNAGAGTACATCAT	957
QY	294	GlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr	313
Db	958	GGACCTGCCACCGAGGCAGCCGATGATCTTTGTAACCTGACGGTGATAAGGAACAAATCACA	1017
QY	314	LeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMet	333
Db	1018	CTTTATGGCCTGAAGAAGGCCAGTCATGGATTGCTCGACCCTCTAAGGGACCCATCATG	1077
QY	334	LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro	353
Db	1078	CTTGGAAGTGTGCTTCTCTTGTCATCTCGTCATGGGAGCATGGTGAACAGAGTGTACCC	1137
QY	354	LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly	373
Db	1138	CTTATGGATCCGATTGTGACACTTTTGTGGTAGTGTCCATGAGAATATGCCTCAAGCTGGA	1197
QY	374	GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln	393
Db	1198	GGAAGTATGAGGAGCACATTGTTTCCAAACTTTGGAAGTATGTTGAGTGTACAGATCAG	1257
QY	394	HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla	413
Db	1258	CATGCCAAAANAATGAGCAGTGGGATGAAGAGAAATCTTCATAGGGATGACGAGGAGTACGCA	1317
QY	414	SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln	433
Db	1318	TCTGATGGTGCAGGAGGTGACTATGAGGACAAATCTCCATAGCCCATTGCTGCCAGGCAG	1377
QY	434	AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer	453
Db	1378	GCAACAGGTGCGGAAGGAAGGACATTGTGCACCATGGTCACCGTGAAGTGTCTTGAGC	1437
QY	454	MetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGly	473
Db	1438	ATGAGAAGGCAAAGCCTCTTTAGGGAGGGTGGAGATGGTGTGAGCAGCACTGATATCGGT	1497
QY	474	GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu	493
Db	1498	GGGGGATGGCAGCTTGCTTGGAAATGGTCAGAGAAGGAAGGTGAGAAATGGTAGAAAGGAA	1557
QY	494	GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer	513
Db	1558	GGTGGTTTCAAAAAGAGTCTACTTGCACCAAGAGGGAGTTCTCTGGCTCAAGAAGGGGCTCA	1617
QY	514	IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla	533
Db	1618	ATTGTTTTCACCTCCCGGTGGTGGCGATGTTCTTGAGGGGTAGTGAGTTGTACATGCTGCT	1677
QY	534	AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp	553
Db	1678	GCTTTAGTAAGTCAGTCAGCACCTTTTCTCAAAGGGTCTTGTCTGAACCAACGTCATGTCAGAT	1737
QY	554	AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe	573
Db	1738	GCTGCCATGGTTACCCCATCTGAGGTAGCTGCCAAAGGTTTCCAGTTGGANAAGATTGTTTT	1797
QY	574	GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe	593
Db	1798	GAACCTGGAGTGAGGCGTGCCCTGTAGTCCGTTGGAATTTCAGATCCTTCAACAGTTTT	1857
QY	594	AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla	613
Db	1858	GCTGGAATAAACGGTGTCTGTACTATATACCCCAAAATTTCTTGAGCAAGCTGGTGTGGCA	1917
QY	614	ValIleLeuSerLysPheGlyLeuSerSerAlaSerIleLeuIleSerSerLeu	633
Db	1918	GTTATTCTTCCAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTC	1977
QY	634	ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly	653



```
Db      1978  ||||| ACTACCTTACTAATGCTTCCTTGCATTGGCTTTGCCATGCTGCTTATGGATCTTTCCGGA 2037
Qy      654  ArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673
Db      2038  AGAAGGTTTTTGTGCTAGGCACAATTCCAATCTGTATAGCATCTCTAGTTATCTCGTT 2097
Qy      674  ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693
Db      2098  GTGTCCAATCTAATTGATTGGGTACACTAGCCCATGCTTTGCTCTCCACCATCAGTGTT 2157
Qy      694  IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713
Db      2158  ATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGGTCCCATCCCCAACATTTTATGTGCA 2217
Qy      714  GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTyr 733
Db      2218  GAGATCTTTCCAACCAGGGTTCGTGGCCTCTGTATTGCCATTTGTGCTTTTACATTCTGG 2277
Qy      734  IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753
Db      2278  ATCGGAGATATCATCGTCACCTACAGCCTTCTCTGTGATGCTGAATGCTATTGGACTGGCG 2337
Qy      754  GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773
Db      2338  GGTGTTTTCAGCATATATGCAGTCGTATGCTTGATTTCTCTTGTGTTCTCTCTTAAG 2397
Qy      774  ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793
Db      2398  GTCCCTTGAGACAAAGGGATGCCCCCTTGAGGTTATTACCGAATTTCTTTGCAGTTGGTGCG 2457
Qy      794  LysGlnAlaAlaAlaLysAla 800
Db      2458  AAGCAAGCGGCTGCAAAAGCC 2478
```

RESULT 4

```
US-10-051-909-1
; Sequence 1, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
```

; LENGTH: 2824

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (29)

; NAME/KEY: unsure

; LOCATION: (622)

; NAME/KEY: unsure

; LOCATION: (636)

; NAME/KEY: unsure

; LOCATION: (638)

; NAME/KEY: unsure

; LOCATION: (669)

; NAME/KEY: unsure

; LOCATION: (771)

; NAME/KEY: unsure

; LOCATION: (822)

; NAME/KEY: unsure

```
; LOCATION: (856)
; NAME/KEY: unsure
; LOCATION: (889)
; NAME/KEY: unsure
; LOCATION: (896)
; NAME/KEY: unsure
; LOCATION: (944)
US-10-051-909-1
Alignment Scores:
Pred. No.: 0
Score: 3517.00
Percent Similarity: 95.18%
Best Local Similarity: 91.57%
Query Match: 86.31%
DB: 13
Gaps: 2
US-10-051-909-32 (1-800) x US-10-051-909-1 (1-2824)
Qy      56  MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75
Db      238  ATGGGGGGCGCCGTGATGGTCGCCATCGCGCCTCTATCGGCAACTTGTCTGCAGGGCTGG 297
Qy      76  AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95
Db      298  GACAATGCGACAATTGCTGGAGCCGTCCTGTACATAAAGAGGAATTCAACCTGCAGAGC 357
Qy      96  GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115
Db      358  GAGCCTCTGATCGAAGCCCTCATCGTCGCCATGTTCTCTCATTTGGGGCAACAGTCATCACA 417
Qy      116  ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135
Db      418  ACATCTCCGGGGCCCAAGGGCTGACTGCGTTGGTAGGAGGCCCATGCTGGTCGCCCTCGGCT 477
Qy      136  IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155
Db      478  GTCCTCTACTTTCGTAGTGGGCTGGTGATGCTTTGGGGCCCAATTGTGTACATCTTGCTC 537
Qy      156  LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175
Db      538  CTGCAAGGCTCATTTGATGGGTTCCGGTATCGGTTTGGCGGTACACTTGTCTCTCTAC 597
Qy      176  IleSerGluIleAlaProSerGluIleArg--GlyLeuLeuAsnThrLeuProGlnPhe 194
Db      598  ATCTCCGAAACTGCACCCGCACAGANATTCTTGGGGCTGNTNGAACACGTTGCCGAGTTC 657
Qy      195  SerGly---SerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSer 213
Db      658  ATTGGGGTCAGNGGAGGATGTTCTCTCTACTGTCATGGTGTGTTGGGATGTCCTCATG 717
Qy      214  ProSerProAspTyrArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePhe 233
Db      718  CCCAAACCTGATTGGAGGCTCATGCTTGGAGTTCTGTCGATCCCCTCACTTATNTACTTT 777
Qy      234  GlyLeuThrIlePheTyrLeuProGluSerProArgTyrLeuValSerLysGlyArgMet 253
Db      778  GGACTGACTGTCTTCTACTTGCCTGAATCACCAAGGTGGCTTGTNAGCAAAAGGAGATG 837
Qy      254  AlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeu 273
Db      838  GCGAGGCGAAGAGAGTGTGCAAGGCTGCGGGGAAGAGAGATGTCTCANGGGAGANG 897
Qy      274  SerLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGlyTyrIleIle 293
Db      898  GCTCTTCTAGTTGAAGGTTTGGGGGTCGGTAAAGATACACGTAITTTNAGAGTACATCAT 957
Qy      294  GlyProAlaThrGluAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThr 313
Db      958  GGACCTGCCACCCGAGGCGCGATGATCTTTGTAAGGAGGAGGAGGAGGAGGAGGAG 1017
Qy      314  LeuTyrGlyProGluGlyGlnSerTyrIleAlaArgProSerLysGlyProIleMet 333
```

Db 1018 CTTTATGGGCCCTGAAGAAGCCAGTCATGGATTGCTCGACCTTCTAAGGGACCCCATCATG 1077

Qy 334 LeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValPro 353

Db 1078 CTTGGAAAGTGCTTTCTCTTGCACTCTCGTCATGGGAGCATGGTGAACACAGAGTGATCCC 1137

Qy 354 LeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly 373

Db 1138 CTTATGGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGATATGCCTCAAGCTGGA 1197

Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393

Db 1198 GGAAGTATGAGGAGCACATTGTTTCCAAACTTTTGGAAAGTATGTTCAAGTGTCAAGATCAG 1257

Qy 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413

Db 1258 CATGCCAAAAATGAGCAGTGGGATGAAGAGAATCTTCCATAGGGATGACGAGGATACGCA 1317

Qy 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433

Db 1318 TCTGTAGTGCAGGAGGTGACTATGAGGACAATCTCCATAGCCCATTTGCTGTCCAGGCAG 1377

Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSer 453

Db 1378 GCAACAGGTGCGGAAGGGAAGGACATGTGCACCATGGTCAACCGTGAAGTGCTTTTGAGC 1437

Qy 454 MetArgArgGlnThrLeuLeuGlyGlyGlyAspGlyGlyValSerSerThrAspIleGly 473

Db 1438 ATGAGAAGGCAAAAGCCTCTTAGGGAGGGTGGAGATGGTGTGACGACACTGATATCGGT 1497

Qy 474 GlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGlu 493

Db 1498 GGGGGATGGCAGCTTGCTTGGAAATGGTCAGAGAAGGAAGGTGAGAATGGTAGAAAGGAA 1557

Qy 494 GlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySer 513

Db 1558 GGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGGAGTTCCTGGCTCAAGAGGGGCTCA 1617

Qy 514 IleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533

Db 1618 ATTTGTTTCACTTCCCGGTGGTGGCGATGTTCTTGAGGGTAGTGAGTTGTACATGCTGCT 1677

Qy 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553

Db 1678 GCTTTAGTAAGTCAGTCAGCACTTTTCTCAAAGGGTCTTGCTGAACACGACATGTCAGAT 1737

Qy 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPhe 573

Db 1738 GCTGCCATGGTTCACCCATCTGAGGTAGTGCCAAAGGTTACAGTTGGAAGATTTGTTT 1797

Qy 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593

Db 1798 GAACCTGGAGTGAGGCGTGCCCTGTTAGTCGGTGTGGAATTCAGATCCTTCAACAGTTT 1857

Qy 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613

Db 1858 GCTGGAAATAACGGTGTCTGTACTATACCCCAAAATTTCTTGACCAAGCTGGTGTGGCA 1917

Qy 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633

Db 1918 GTTATTCTTCCAAAATTTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTC 1977

Qy 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653

Db 1978 ACTACCTTACTAATGCTTCTCTGCAATTGGCTTTGCCATGCTGCTTATGGATCTTTCCGGA 2037

Qy 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673

Db 2038 AGAAGGTTTTTGCTGCTAGGCACAATTCCAATCTTGATAGCATCTTAGTTATCCTGGTT 2097

Qy 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693

Db 2098 GTGTCCAATCTAATTGATTTGGGTACACTAGCCCCATGCTTTGTCTCTCCACCATCAGTGT 2157

Qy 694 IleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713

Db 2158 ATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGGTCCCATCCCCAACATTTTATGTGCA 2217

Qy 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733

Db 2218 GAGATCTTTCCAACCAGGGTTCGTGGCCTCTGTATTGCCATTTGTGCCTTTACATTCGG 2277

Qy 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753

Db 2278 ATCGGAGATATCATCGTCACCTACAGCCTTCCTGTGATGCTGAATGCTATTGGACTGGCG 2337

Qy 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773

Db 2338 GGTGTTTTTCAGCATATATGCAGTCGTATGCTTGATTTCCTTTGTGTTCGTCTTCCTTAAG 2397

Qy 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793

Db 2398 GTCCCTGAGACAAAGGGGATGCCCTTGAGGTTATTACCGAATTCCTTGCAGTTGGTGCG 2457

Qy 794 LysGlnAlaAlaAlaLysAla 800

Db 2458 AAGCAAGCGGCTGCAAAAGCC 2478

RESULT 5

US-10-425-114-32961

; Sequence 32961, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 32961

; LENGTH: 2031

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: UC-ZMFLMO17045C09\_FLI

US-10-425-114-32961

Alignment Scores:

Pred. No.: 3.75e-274 Length: 2031

Score: 2869.00 Matches: 561

Percent Similarity: 99.65% Conservative: 1

Best Local Similarity: 99.47% Mismatches: 2

Query Match: 70.40% Indels: 0

DB: 17 Gaps: 0

US-10-051-909-32 (1-800) x US-10-425-114-32961 (1-2031)

Qy 237 IlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGluAla 256

Db 2 ATATTATCTTCTCTGAATCTCCAAGATGGCTCGTTAGCAAAAGTTCGGATGGCAGAGGCA 61

Qy 257 LysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeuLeu 276

Db 62 AAAAAAGGTGTGCAAAAGTTACGGGGGAAAGACGATGTCTCAGTGAATTTGTCCTTCTT 121

Qy 277 LeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyProAla 296

Db 122 CTCGAAGGGTTGGAGGTTGGAGGAGACACTTCCATTGAAGAGTACATCATTTGGACCTGCC 181

Qy 297 ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGly 316

Db 182 ACCGAGGCAGCCGATGATCTTGTTACTGACGGTGATAAGGAACAAATCACACTTTATGGG 241  
Qy 317 ProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGlySer 336  
Db 242 CCTGAAGAAGGCCAGTCATGGATTGCTCGACCTTCCAAGGGACCCAGCATGCTTGGAAGT 301  
Qy 337 ValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAsp 356  
Db 302 GTGCTTCTCTTGTCATCTCGTCATGGGAGCATGGTGAACACAGAGTGACCCCTTATGGAT 361  
Qy 357 ProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMet 376  
Db 362 CCGATTGTGACACTTTTGGTAGTGTCCAATGAGAAATATGCCTCAAGCTGGAGGAATATG 421  
Qy 377 ArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLys 396  
Db 422 AGGAGCACATTGTTTCCAAACTTTGGAAGTATGTTCAAGTGTCCAGATCAGCATGCCAAA 481  
Qy 397 AsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTrpAlaSerAspGly 416  
Db 482 AATGAGCAGTGGGATGAAGAGAAATCTTCATAGGATGACGAGGAGTACGCATCTGATGGT 541  
Qy 417 AlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGly 436  
Db 542 GCAGGAGGTGACTATGAGACAATCTCCATAGCCCATGCTGTCTCCAGGAGGCAACAGGT 601  
Qy 437 AlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaAlaLeuSerMetArgArg 456  
Db 602' GCGGAAGGAAGGACATTGTGCACCATGGTCCCGTGAAGTGTCTTGGAGCATGAGAAGG 661  
Qy 457 GlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrp 476  
Db 662 CAAAGCCTCTTAGGGAGGGTGGAGATGGTGTGACGAGCACTGATATCGTGGGGGATGG 721  
Qy 477 GlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPhe 496  
Db 722 CAGCTTGCTTGAAATGGTCAAGAGGAAGGTGAGAATGGTAGAAGGAAGGTGGTTTC 781  
Qy 497 LysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSer 516  
Db 782 AAAAGAGTCTACTTGCACCAAGAGGGAGTTCTCGCTCAAGAGGGGCTCAATTGTTTCA 841  
Qy 517 LeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuVal 536  
Db 842 CTTCCCGGTGGTGGCGATGTTCTTGAGGGTAGTGAGTTGTACATGCTGCTTTAGTA 901  
Qy 537 SerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMet 556  
Db 902 AGTCAGTCAGCACTTTCTCAAAGGGTCTTGCTGAACCAACGATGTACATGCTGCCATG 961  
Qy 557 ValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGly 576  
Db 962 GTTCACCCATCTGAGGTAGTGCCAAAGGTTCAAGTTGGAAAGATTTGTTGAACCTGGA 1021  
Qy 577 ValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIle 596  
Db 1022 GTGAGGCGTGCCCTGTTAGTCGGTGTGGAAATTCAGATCCTTCAACAGTTTGTGGAATA 1081  
Qy 597 AsnGlyValLeuTyrTrpThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeu 616  
Db 1082 AACGGTGTTCTGTACTATATACCCACAAAATTTCTTGAGCAAGCTGGTGTGGCAGTTATTCTT 1141  
Qy 617 SerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrLeu 636  
Db 1142 TCCAAATTGGTCTCAGCTCGGCATCAGCATCCATCTTGATCAGTTCTCTCACTACCTTA 1201  
Qy 637 LeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPhe 656  
Db 1202 CTAATGCTTCTTGTCATTGGCTTTGGCATGTGCTGCTTATGGATCTTTCCGGAAGAAGGTTT 1261  
Qy 657 LeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsn 676

Db 1262 TTGCTGCTAGGCACAATTCCAATCTTGATAGCATCTCTAGTTATCTTGTTGTGTCCAAT 1321  
Qy 677 LeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyr 696  
Db 1322 CTAATTGATTGGGTACACTAGCCCATGCTTTTGCTCTCCACCGTCAGTGTATCGTCTAC 1381  
Qy 697 PheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePhe 716  
Db 1382 TTCTGCTGCTTCGTTATGGGATTTGGTCCCATCCCCAACATTTTATGTGCAGAGATCTTT 1441  
Qy 717 ProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAsp 736  
Db 1442 CCAACCAGGGTTCGTGGCCTCTGTATTGCCATTTGTGCCTTTACATTTCTGGATCGGAGAT 1501  
Qy 737 IleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPhe 756  
Db 1502 ATCATCGTCACCTACAGCCTTCCTGTGATGCTGAATGCTATTGGACTGGCGGGTGTTC 1561  
Qy 757 SerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGlu 776  
Db 1562 AGCATATATGCAATCGATGCTTGTGATTTCCTTTGTGTCGCTTCCTTAAGGTCCCTGAG 1621  
Qy 777 ThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla 796  
Db 1622 ACAAGGGGATGCCCTTGAGGTTATACCGAATTCCTTGCAGTTGGTGCGAAGCAAGCG 1681  
Qy 797 AlaAlaLysAla 800  
Db 1682 GCTGCAAAAGCC 1693  
RESULT 6  
US-10-437-963-93472/c  
; Sequence 93472, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 93472  
; LENGTH: 2769  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_91853C.1  
US-10-437-963-93472  
Alignment Scores:  
Pred. No.: 1-29e-270 Length: 2769  
Score: 2835.50 Matches: 562  
Percent Similarity: 82.45% Conservative: 77  
Best Local Similarity: 72.52% Mismatches: 120  
Query Match: 69.58% Indels: 16  
DB: 18 Gaps: 7  
US-10-051-909-32 (1-800) x US-10-437-963-93472 (1-2769)  
Qy 30 AlaLeuProGlyProLeuProPro-AlaSerCys-----SerSerGl 43  
Db 2684 GGCCTTCCAGATTCACGCGCCTCTCTTCTTGTAGGGATCCGAAATCTCGGTGGACG 2625  
Qy 43 nGluProValThrSerAspAspIleLeuGluAspLysMetSerGlyAlaValLeuValAl 63



Db 2624 AGAGACTTGGTGAAGAT-----TCGCCGCCCATGGCGGGCCCGTCTGGTCGC 2574

Qy 63 aileValAlaSerIleGlyAsnLeuLeuGlnGlyTrpAspAsnAlaThrIleAlaAa1 83

Db 2573 CATCGGCGCTCCATCGGCAACTTGTCTGCAGGCTGGGATAATGCAACCATTCAGGTGC 2514

Qy 83 aValLeuTyriIleLysGluPheGlnLeuGlnAsnGluProThrValGluGlyLeu11 103

Db 2513 GGTACTGTACATCAAGAAGGAATTCAACTTGCAGCGGAGCCCTTATCGAAGGCCTGAT 2454

Qy 103 eValSerMetSerLeuIleGlyAlaThrIleValThrPheSerGlyProLeuSerAs 123

Db 2453 CGTGGCCATGTCGCTCATTTGGGCGACGATCATCAGCAGCTTCTCTGGAGCAGTGGCTGA 2394

Qy 123 pSerIleGlyArgArgProMetLeuIleLeuSerSerIleLeuTyriPhePheSerGlyLe 143

Db 2393 TTCTTTTGGTAGCGGCCCATGCTGATCGCGTCGGCTGTCTCTACTTTTGTAGTGGCT 2334

Qy 143 uIleMetLeuTrpSerProAsnValTyriValLeuLeuAlaArgPheValAspGlyPh 163

Db 2333 AGTGATGCTTTGGCGCCAAATGTGTATGTGTGTCTCTTGGCGAGGCTCATTGACGGGTT 2274

Qy 163 eGlyIleGlyLeuAlaValThrLeuValProLeuTyriIleSerGluIleAlaProSerG1 183

Db 2273 CGGGATCGGTTTGGCTGTACGGCTTGTACCATGTATCATCTCTGAGACTGCCCCGACGGA 2214

Qy 183 uIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeuSe 203

Db 2213 CATCAGAGGACTGCTAAACACAGCTGCCGCGAGTTCAGTGGGTCTGGAGGGATGTTCTTTC 2154

Qy 203 rTyriCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuG1 223

Db 2153 ATACTGCATGGTATTGGCATGTCCCTCATGCCACAGCCAGATTTGGAGGATCATGCTTGG 2094

Qy 223 yValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyriLeuProGluSe 243

Db 2093 CGTTCTATCAATACCATCACTTATATACTTTGCATTGACCATCTTTTACTTACCTGAATC 2034

Qy 243 rProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLe 263

Db 2033 GCCGAGGTGGCTCGTGAGCAAGGAAGAAATGGCTGAGGCCAAGCGTGTGTGCAAGGCCT 1974

Qy 263 uArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuValG1 283

Db 1973 GCGTGAAGAGAGAAGATGTTTCAGGAGAAATGGCCCTTCTCGTTGAAGGTCTGGGGTTGG 1914

Qy 283 yGlyAspThrSerIleGluGluTyriIleIleGlyProAlaThrGluAlaAaAspLe 303

Db 1913 GAAAGACACAAAAATTGAGGAATACATAATTGGACCTGATGATGAGCTTGCTGATGAAGG 1854

Qy 303 uValThrAspGlyAspLysGluGlnIleThrLeuTyriGlyProGluGluGlyGlnSerTr 323

Db 1853 GCTGGCTCCA---GATCCAGAGAAGATCAAACTGTATGGTCTCTGAAGAGGCTTATCGTG 1797

Qy 323 pIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerAr 343

Db 1796 GGTGGCCCGTCTCTTCACGGGCAAGTGCACCTTGGAAAGTGCAATTAGGTCTCATCTCTCG 1737

Qy 343 gHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheG1 363

Db 1736 TCATGGTAGTATGTCAGTCAGGGTAAGCCCTTGTGGATCCTGTTGCACCCCTTTTGG 1677

Qy 363 ySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAs 383

Db 1676 AAGTGTCCATGAGAAGATGCCTGAGATAATGGGAAGCATGCGGAGCACAATTGTTTCCTAA 1617

Qy 383 nPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluG1 403

Db 1616 CTTTGGCAGCATGTTTAGTGTGGCGGAACAGCAGCAAGCTAAAGGTGATTGGGATGCTGA 1557

Qy 403 uAsnLeuHisArgAspAspGluGluTyriAlaSerAspGlyAlaGlyGlyAspTyriGluAs 423

Db 1556 GAGT---CAACGGGAGGGTGAAGATTATGGATCAGACCATGGTGGGGATGACATTGAAGA 1500

Qy 423 pAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle-- 442

Db 1499 TAGCCTCCAAAGCCCACTTATTTCTCGTCAAGCGACAGCGTGGAAAGGAGATCGC 1440

Qy 443 -ValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyG1 462

Db 1439 TGCACCTCATGGCAGTATAATGGGTGCTGTGGGA-----AGAAGTAGTAGTCTCATGCA 1386

Qy 462 uGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTr 482

Db 1385 GGGCGGGAGGCAGTAAGCAGCATGGSCATTGGTGGGGATGGCAGTTGGCTTGGAAATG 1326

Qy 482 pSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyriLeuH1 502

Db 1325 GACTGAGAGAGAAAGTGCAGATGGCGAAAAGAGGTGGCTTCCAACGTATCTACTTGCA 1266

Qy 502 sGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAa 522

Db 1265 TGAAGAGGGTGTGACAGGTGATCGCAGGGGCTCTATACTGTCTATTCCT---GGAGGTGA 1209

Qy 522 pValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPh 542

Db 1208 TGTTCTCTCCTGGTGGTGAGTTCGTCAGGCAGCTGCTCTTGTAGCCAAACCTGCTCTTTA 1149

Qy 542 eSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluVa 562

Db 1148 CTCTAAGGAATTGATGGAGCAACGCCTTGCTGGCCCTGCTATGTTGCATCCATCTCAGGC 1089

Qy 562 lAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLe 582

Db 1088 AGTTGCTAAAGGTCCAAAATGGGCGACATTATTGGAACCTGGAGTGAAGCATGCTCTGTT 1029

Qy 582 uValGlyValGlyIleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyriTy 602

Db 1028 TGTTGGCATAGGGATACAAATCCTGCAACAGTTCGTTGGCATTAAATGGAGTTCGTACTA 969

Qy 602 rThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSe 622

Db 968 CACTCCACAAAATTCTTGAGCAAGCTGGTGTGGTGTCTTCTTTCGCAAAACATTTGGACTTAG 909

Qy 622 rSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCys11 642

Db 908 CTCCTCATCTGCATCTATTCTTTATAGCGGACTGACAACCTTGTGTGATGCTTCCCAGCAT 849

Qy 642 eGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThr11 662

Db 848 TGGTATTGCTATGAGGCTCATGGATATGTCTGGAAGAAGGTTTCTTCTCTTTCGCAACAAT 789

Qy 662 eProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyTh 682

Db 788 CCCTATCTGTATAGTACACTAGCTATCTTGATTCTGGTCAATATTCTGGATGTGGGGAC 729

Qy 682 rLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyriPheCysCysPheValMe 702

Db 728 CATGGTTTCATGCCTCACCTGTCCACAGTCAGTGTCTACTCTACTTCTGTCTTCTTGTTCAT 669

Qy 702 tGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgG1 722

Db 668 GGGGTTTCGGGCGCTATTCCAAACATTCCTCTGTGCAGAGATTTTCCCGACCACCGTTTCGTGG 609

Qy 722 yLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyriSe 742

Db 608 CATCTGCATAGCCCATCTGTGCCCCTAACATTCCTGGATCGGTGATATCATTTGTGACATACAC 549

Qy 742 rLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyriAlaValva 762

Db 548 CCTCCCCGTGATGCTCAACGCCATTGGACTCGCTGGAGTGTTTGGAAATCTACGCAGTGGT 489

Qy 762 lCysLeuIleSerPheValPheLeuLysValProGluThrLysGlyMetProle 782

Db 488 CTGCATACTGGCTTTCCTGTTTGTCTTTCATGAAGGTGCCGGAGACAAAGGGCATGCCTCT 429

QY 782 uGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla 796  
Db 428 TGAAGTCATCACCGAGTCTCTCTGTCTGGAGCAAGCAGGCC 386

RESULT 7  
US-10-425-115-70742  
; Sequence 70742, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 70742  
; LENGTH: 4330  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_164513C.1  
US-10-425-115-70742

Alignment Scores:  
Pred. No.: 1.2e-265 Length: 4330  
Score: 2788.50 Matches: 538  
Percent Similarity: 84.23% Conservative: 87  
Best Local Similarity: 72.51% Mismatches: 110  
Query Match: 68.43% Indels: 7  
DB: 18 Gaps: 6

US-10-051-909-32 (1-800) x US-10-425-115-70742 (1-4330)

QY 56 MetSerGlyAlaValLeuValAlaValIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75  
Db 769 ATGGGGGCGCGTGTATGTCGCCATCGCGCCTCTATCGCAACTTGTGCGGGCTGG 828  
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95  
Db 829 GACAATGCGACAATTGCTGGAGCCGTCCTGTACATAAAGAGGAATTCAACCTGCAGAGC 888  
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
Db 889 GAGCCTCTGATCAGAGGCGCTCATCGTCGCCATGTCCCTCATTTGGGGCAACAGTCATCACG 948  
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
Db 949 ACGTTCCTCCGGGCGAGCGCGGACTGCGTGGTAGGAGGCCCATGCTGTCGCTCGGCT 1008  
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155  
Db 1009 GTCCTCTACTTCGTCAAGTGGGCTGGTGTGATGCTCTGGCGCCCAAGTGTGTATCTTGTCTC 1068  
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
Db 1069 CTCGCAAGGCTCATTTGATGGGTTCCGGTATCGGTTTGGGGTCACTTGTTCCTCTCTAC 1128  
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
Db 1129 ATCTCCGAGACTCGCGCGACAGACATTCGTGGGCTGTTGAACACGTTGCCGAGTTTCAGT 1188  
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
Db 1189 GGGTCAGGAGGGATGTTCCTCTCTCTACTGCAATGGTGTGGGATGTCCTCATGCCCCAAA 1248  
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
Db 1249 CCTGATTGGAGGCTCATGCTGGAGTTCTGTCGATCCCGTCACTTATTACTTTGGACTG 1308

QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
Db 1309 ACTGTCTTCTACTTGCCTGAATCACCAAGGTGGCTTGTGAGCAAGGAAGGATGCGGAG 1368  
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275  
Db 1369 GCGAAGAGAGTGTTCGAAAGGCTCGGGGAAGAGAGATGTCTCAGGGGAGATGGCTCTT 1428  
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
Db 1429 CTAGTTGAAGGTTTGGGGTCCGTTAAAGATACACGTATTGAAGAATACATAATTGGTCCC 1488  
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
Db 1489 GATGATGAACCTGCTGATGAAGGCTGGCTCCA--GATCCAGAGAAGATCAAACTATAT 1545  
QY 316 GlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
Db 1546 GGACCTGAAGAGGCTATCTTGGGTTGCCCGACCTGTTCCGGGACAAAGTGTCTTGGGA 1605  
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMet---ValAsnGlnSerValProLeu 354  
Db 1606 AGCGGTTAGGTCATCTCTCGTCATGGGAGTATGGCGGTAGTCAGGGTAAGCCCTC 1665  
QY 355 MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly 374  
Db 1666 GTGGATCCTATGCTCACTCTTTTCGGAAGTGTTCATGAAAAGATGCCTGAGATCATGGGG 1725  
QY 375 SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis 394  
Db 1726 AGCATGAGGAGGACATGTTTCCCAACTTTGGCAGCATGTTTAGTGTTCGCCAGCAGCAG 1785  
QY 395 AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer 414  
Db 1786 CAGGTGAAAGCTGACTGGGACGCCGAGAGT---CAAAGGGAAGGTGAAGATTATGCTTCG 1842  
QY 415 AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434  
Db 1843 GATCATGTTGGCGATGACATTTGAGGATAACCTCCAAAGCCCACTTATTTCGTGAGGCA 1902  
QY 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454  
Db 1903 ACAAGTGTGAAGGAAGAGAGATCGCT-----GCACCTCATGTTAGCATATTGGGTGCT 1956  
QY 455 ArgArgGlnThrLeuLeuGlyGlyGlyAspGlyValSerSerThrAspIleGlyGly 474  
Db 1957 GTGGAAAGGAGCAGTAGCTTGACGGGAGGGAGGAGCAGTAAGCAGCATGGGCATTTGCCGA 2016  
QY 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494  
Db 2017 GGATGGCAGTTGGCGTGGAAATGGACCGAGAGAGAGGGCGAAGATGGGCAAAAGGAAGGT 2076  
QY 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514  
Db 2077 GGCCTCCAGCGTATTTACTTGCATGAGGAGGGCGGTACAAGGCAAC---AGGGGTTCTATA 2133  
QY 515 ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 534  
Db 2134 TTGTCAATACCA--GGCGGGATGTTCTCTCTGTTGGTGGTGGTTCATCCAGGCTGCAGCT 2190  
QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554  
Db 2191 CTTGTGAGCCAAACAGCTCTTTACTCTAAGGAACCTGCTGGAGCAACCTGCTGCTGCTCCT 2250  
QY 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574  
Db 2251 GCGATGATGCATCCATCTGAAGCAGTTACTAAAGGTCCAAGATGGGCCGACCTATTGTAG 2310  
QY 575 ProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAla 594  
Db 2311 CCTGGGGTGAAGCATGCACCTGTTGTTGGCATAGGAATACAGATCCTGCAACAGTTTGCT 2370  
QY 595 GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614

Db 2371 GGCATCAACGGCGTTCTCTACTACACTCCTCAAATCTTGAGCAAGCAGCGTCGGTGTT 2430  
QY 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634  
Db 2431 CTTCTGTCGAACCTCGGCCCTTAACGCTTCTTCGGCATCAATCCTCATAGCGCCTTGACG 2490  
QY 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654  
Db 2491 ACCTTACTGATGTCCTCCAAGCATCGGCATTGCGATGAGGCTCATGGATATGTCGGAAG 2550  
QY 655 ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674  
Db 2551 AGGTTTCTCCTCGCGACGATCCCAGTCTTAATAGTCGCGCTACTCGTCTCGTGGTG 2610  
QY 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694  
Db 2611 TCCAACATCGTGACGTGGGGACGTGGCGCACGGCGGCTCTCCACGGCCAGCGTCATA 2670  
QY 695 ValTyrPheCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714  
Db 2671 GTCTACTTCTGCTTCTCGTCAATGGGGTTTCGGGCCCGTCCCCAACATCCTCTGCGCAGAG 2730  
QY 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734  
Db 2731 ATCTTCCCCACACACGGTCCGCGGTGCTCGCATCGCCATCTGCGCCCTGGCCTTCTGGGTC 2790  
QY 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754  
Db 2791 GGTGACATCATCGTGACGTACACTCTCCCGGTGATGCTGAACGTCTCGGGCTCGCCGGC 2850  
QY 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774  
Db 2851 GTCTTTGGGGTGACGCCGTGCTGTGCGTCCTAGCCCTCGCGTTCGTTCGTCAAGGTG 2910  
QY 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794  
Db 2911 CCCGAGACGAAGGGCATGCCTCTCGAGGTGTCATCCCGAGTTCTTCTCCGTTGGGGCAAAG 2970  
QY 795 GlnAla 796  
Db 2971 CAAGCC 2976

RESULT 8  
US-10-051-902-7  
; Sequence 7, Application US/10051902  
; Publication No. US20020178468A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/10/051,902  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/291,922  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-14  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 2601  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-10-051-902-7  
Alignment Scores:  
Pred. No.: 1.28e-254 Length: 2601  
Score: 2674.00 Matches: 522  
Percent Similarity: 82.20% Conservative: 92  
Best Local Similarity: 69.88% Mismatches: 115  
Query Match: 65.62% Indels: 18  
DB: 13 Gaps: 8

US-10-051-909-32 (1-800) x US-10-051-902-7 (1-2601)  
QY 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75  
Db 175 ATGAAAGTGCCGTCCTTGTTGCTATTGCGCGCTTCCATTGGTAATTTCTCTCAAGGATGG 234  
QY 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95  
Db 235 GATAATGCTACCATCGCCGGGCTAATGGTTACATTAAAGAAAGACCTTGCTTTGGGAACA 294  
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
Db 295 -----ACTATGGAAAGGCTTGCTGGTGGCATGTCCCTGATTGGAGCAACGGTAATCACC 348  
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
Db 349 ACATGCTCTGGTCCTATAGCGGATGGCTCGGTCCGCCACCCATGATGATAATCTCATCT 408  
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155  
Db 409 GTGCTCTATTCTTGGTGGTTTGGTGATGCTGTGGTCCCCAAATGTGTATGTGTGTGC 468  
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
Db 469 TTGGCAGGCTACTTGTATGGATTGGGATTGGCCTTGCTGTGACTCTTGTCCCGTCTAT 528  
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
Db 529 ATATCTGAACCGGCCGCTCTGAATAAGGGGTCGTTGTAATACGTTCTCTCAGTTTCAGT 588  
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
Db 589 GGCTCTGGAGGAATGTTTTTGTCTGTACTGTATGGTTTTTGGCATGTCTATGAGTCCCGCG 648  
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
Db 649 CCTAGCTGGAGGCTCATGCTTGGGGTCTGTCTATTCTCTCTCTCTTGTATTGTGATTG 708  
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
Db 709 ACCATTTTTTCTTGGCCGAGTCTCCTCGGTGGTGGTCAGCAAGGAAGGATGCTCGAG 768  
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu 275  
Db 769 GCTAAGAAGGTGCTCCAAGATTGCGCGGAAGGGAGGATGTTGTCAGCGAGATGGCATTG 828  
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
Db 829 CTGTTGAAGGTCTCGGATTGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT 888  
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
Db 889 GCTGACGATGTGGCTGATGGTCATGAACATGCAACAGAGAAGATAAAATTCGATTATAT 948  
QY 316 GlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
Db 949 GGATCCCAAGCAGGCCCTTTCITGGTTATCAAAACCTGTCACTGGACAGAGTTCTATTGGC 1008  
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355  
Db 1009 -----CTTGGCTCACCATGGAGCATCATCAACCAAGCATGCCCTCATG 1056  
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373  
Db 1057 GATCCTCTGGTGACACTGTTTGGTAGCATTCATGAGAAAGTCCCCGAGACAGGACAAGA 1116  
QY 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393  
Db 1117 GGAAGCATCGAAGCACCTCTGTTCCAAATTTTGAAGCATGTTTCAGCACTGCTGAGCCG 1176  
QY 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413



Db 1177 CATGCTAAATTTGAACAATGGGATGAAGAAAGCTTACAAAGGGAACGTGAGGACTACATG 1236

Qy 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433

Db 1237 TCAGATGCAACCCGTCGGGACTCCGATGATAATTTGCACAGTCCTTTAAATCTCACGCCAA 1296

Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisGlyHisArgGlySerAlaLeu--- 452

Db 1297 ACAACAAGCCTTGAA---AAAGACTTACCTCCTCCTCTTCCCATGGCAGTATCCTTGGC 1353

Qy 453 SerMetArgArgGlnThrLeuLeuGlyGluGly---GlyAspGlyValSerSerThrAsp 471

Db 1354 AGCATGAGCGTCACAGTAGTCTCATGCAAGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413

Qy 472 IleGlyGlyGlyTyrGlnLeuAlaTyrLysTyrSerGluLysGluGlyGluAsnGlyArg 491

Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAAATGGACTGATAAA---GGTGAGGATGGAAAA 1470

Qy 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511

Db 1471 CAACAAGGAGGTTTTAAAGGATTTATTTACATGAGGAGGAGTTTTCGCATCTCGTCGT 1530

Qy 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531

Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTTGTCCAG 1572

Qy 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551

Db 1573 GCTGCTGCCTTGGTAAGCCAACCCGCTCTTTACTCCAAGGAGCTTATGTGACACCCA 1632

Qy 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAsp 571

Db 1633 GTTGGCCTGCAATGTTTACCCCATCTGAGACAGCTTCAAAGGGGCCCAAGTTGGAAGCT 1692

Qy 572 LeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591

Db 1693 CTTCCTTGAACCAAGGGTTAAGCATGCAATTGGTTGTGGAGTTGGAATACAAATACTTTCAG 1752

Qy 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly 611

Db 1753 CAGTTTTCAGGGATAAATGGGGTCTATATATTACACCTCAAATCCTTGAAGAGGCCGGT 1812

Qy 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631

Db 1813 GTTGAAGTTCTTCTTTCAGATATAGGCATTTGGCTCAGAGTCGGCATCATTCCTTATCAGT 1872

Qy 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651

Db 1873 GCITTCACAAACCTTCTTGATGCTTCCCTGTATAGCGGTAGCCATGAAGCTCATGGATGTT 1932

Qy 652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671

Db 1933 TCAGCAGAAGGCAGTTGCTACTTACTACAATCCCGTGTGATTGTGTCACTCATATT 1992

Qy 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691

Db 1993 TTGGTCATTTGGAAGCCTGGTAAATTTTGGCAATGTGCGCCCATGCAATCTCAACAGTA 2052

Qy 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711

Db 2053 TGCCTTGTGGTTTATTTCTGCTGCTTTGTGATGGGTATGGACCAATTCCAACATCCTT 2112

Qy 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731

Db 2113 TGCTCAGAGATTTTCCCCACTAGGGTGGTGGCCTCTGCATTGCTATCTGTGTCATTAGTG 2172

Qy 732 PheTyrPheGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751

Db 2173 TTCCTGGATTGGAGACATCATCATCACATACTCGCTGCTGTGATGCTCGCTCTTTAGGA 2232

Qy 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771

Db 2233 CTTGGTGGTGTATTGCGCAATTACGCAGTTGTTTGTGTTTCATCTCGTGGATATTGTGTTT 2292

Qy 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791

Db 2293 TTGAAGGTTCCAGAAACAAAGGGCATGCCCTTTGAAGTCATCTCTGAATTCTTTCTGTT 2352

Qy 792 GlyAlaLysGlnAlaAla 798

Db 2353 GGAGCAAAAGCAGGCTGCTTCT 2373

RESULT 9

US-10-051-909-7

; Sequence 7, Application US/10051909

; Publication No. US20020199217A1

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Helentjaris, Tim

; APPLICANT: Hitz, Bill

; APPLICANT: Kinney, Tony

; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins

; FILE REFERENCE: B01163 US CIP

; CURRENT APPLICATION NUMBER: US/10/051,909

; CURRENT FILING DATE: 2002-01-17

; PRIOR APPLICATION NUMBER: 60/083,044

; PRIOR FILING DATE: April 24, 1998

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 7

; LENGTH: 2601

; TYPE: DNA

; ORGANISM: Glycine max

US-10-051-909-7

Alignment Scores:

Pred. No.: 1.28e-254 Length: 2601

Score: 2674.00 Matches: 522

Percent Similarity: 82.20% Conservative: 92

Best Local Similarity: 69.88% Mismatches: 115

Query Match: 65.62% Indels: 18

DB: 13 Gaps: 8

US-10-051-909-32 (1-800) x US-10-051-909-7 (1-2601)

Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr 75

Db 175 ATGAAAGGTGCGTCCTTGTGTGCTATTGCCCTTCCATTGGTAATTTCTCCAAGGATGG 234

Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95

Db 235 GATAATGCTACCATGCGCGGGGCTAATGGTTACATTAAAGAAACCTTGCTTTGGGAACA 294

Qy 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115

Db 295 -----ACTATGAAAAGGCTTGTGTGGGCATGTCCCTGATTGGAGCAACGGTAATCACC 348

Qy 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135

Db 349 ACATGCTCTGGTCCTATAGCGGATTGGCTCGGTCCGGACCCCATGATGATAATCTCATCT 408

Qy 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu 155

Db 409 GTGCTCTATTTCTTGGGTGGTTTGGTGATGCTGTGGTCCCCCAATGTGTATGTGTGTGC 468

Qy 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175

Db 469 TTGGCGAGGCTACTTGATGGATTGGGATTGGCCTTGTGTGACTCTTGTCCCGTCTAT 528

Qy 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195

Db 529 ATATCTGAAACGGCGCGCTCTGAAATAAAGGGGTCGTTGAATACGCTTCTTCAGTTCAGT 588

Qy 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215

Db 589 GGCTCTGGAGGAATGTTTTTGTGCTACTGTATGGTGTCTTGGCATGTCAATGAGTCCCGCG 648

Qy 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235

Db 649 CCTAGCTGGAGGCTCATGCTTGGGTTCTGTCTATTCCTTCTCTCTGTATTTGCAATG 708

Qy 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255

Db 709 ACCATTTTTTCTTGTCCCGAGTCTCCTCGGTGGTGTGTCAGCAAGGAGGATGCTCGAG 768

Qy 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275

Db 769 GCTAAGAAAGGTGCTCCAAAGATTGCGCGAAGGGAGGATGTGTACGGCGAGATGGCATTG 828

Qy 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295

Db 829 CTGGTTGAAGTCTCGGGATTCGGGTTGGGGTGATACATCTATCGAAGAGTACATAATTGGCCCT 888

Qy 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315

Db 889 GCTGACGATGTGGCTGATGGTTCATGAACATGCAACAGAGAAAGATAAAATTCGATTATAT 948

Qy 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335

Db 949 GGATCCCAAGCAGGCCCTTTCTTGGTTATCAAAACCTGTCACTGGACAGAGTTCTATTGGC 1008

Qy 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355

Db 1009 -----CTTGGCTCACACCATGGAAGCATCATCAACCAGCATGCCCCCTCATG 1056

Qy 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGly----- 373

Db 1057 GATCCTCTGGTGACACTGTTTGGTAGCACTCATGAGAAGCTCCCCGAGACAGGAGCAAGA 1116

Qy 374 GlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGln 393

Db 1117 GGAAGCATGCGAAGCACTCTGTTTCCAAATTTTGAAGCATGTTTCAGCACTGCTGAGCCG 1176

Qy 394 HisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAla 413

Db 1177 CATGCTAAATTTGAACAATGGGATGAAGAAAGCTTACAAAGGGAACGTGAGGACTACATG 1236

Qy 414 SerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGln 433

Db 1237 TCAGATGCAACCCGTGGGACTCCCGATGATAAATTTGCACAGTCTCTTAATCTCACGCCAA 1296

Qy 434 AlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeu--- 452

Db 1297 ACAACAAGCCTTGAA---AAAGACTTACCTCCTCCTCCTCCTCCTCCTCCTCCTTGGC 1353

Qy 453 SerMetArgArgGlnThrLeuLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 471

Db 1354 AGCATGAGGCGTCACAGTAGTCTCATGCAAGGTCAGGTGAGCAAGGTGGTAGTACAGGT 1413

Qy 472 IleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArg 491

Db 1414 ATTGGTGGTGGCTGGCAACTGGCATGGAATGGACTGATAAA---GGTGAGGATGGAAA 1470

Qy 492 LysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArg 511

Db 1471 CAACAAGGAGGGTTTAAAGGATTTATTACATGAGGAGGGAGTTTCTGCATCTCGTCGT 1530

Qy 512 GlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHis 531

Db 1531 GGATCCATTGTATCGATTCCCGGTGAAGGC-----GAATTTGTCCAG 1572

Qy 532 AlaAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMet 551

Db 1573 GCTGCTGCCTTGGTAAGCCAACCCGCTCTTACTCCAAGGAGCTTATTGATGGACACCCA 1632

Qy 552 SerAspAlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAsp 571

Db 1633 GTTGGGCCCTGCAATGGTTTCAACCATCTGAGACAGCTTCAAAGGGGCCAAGTTGGAAAGCT 1692

Qy 572 LeuPheGluProGlyValArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGln 591

Db 1693 CTTCTTGAACACAGGGTTAAGCATGCATTTGGTTGGAGTTGGAATACAAATACTTTCAG 1752

Qy 592 GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluAlaGly 611

Db 1753 CAGTTTTCAGGGATAAATGGGGTTCTATATTACACACCTCAAATCCTTGAAGAGGCCGCT 1812

Qy 612 ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer 631

Db 1813 GTTGAAGTTCTTCTTTCAGATATAGGCATTTGGCTCAGAGTCGGCATCATTCCTTATCAGT 1872

Qy 632 SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu 651

Db 1873 GCCTTTCACAACCTTCTTGTATGCTTCCCTGTATAGGCGTAGCCATGAAGCTCATGGATGTT 1932

Qy 652 SerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle 671

Db 1933 TCAGGCAGAAGGCAGTTTGCTACTTACTACAATCCCCGTGCTGATTTGTGTCACTCATTTT 1992

Qy 672 LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal 691

Db 1993 TTGGTCATTTGGAAGCCTGGTAAATTTTGGCAATGTGCCCCATGCAGCAATCTCAACAGTA 2052

Qy 692 SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu 711

Db 2053 TGCCTGTGGTTTATTTCTGCTGCTTGTGATGGGTTATGGACCAATTCGAAACATCCTT 2112

Qy 712 CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr 731

Db 2113 TGCTCAGAGATTTTCCCCACTAGGTCGTGGCTCTGCAATGCTATCTGTGCATTAGTG 2172

Qy 732 PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly 751

Db 2173 TTCTGGATTGGAGACATCATCATCATACTACTCGTGCCTGTGATGCTCGGCTCTTTAGGA 2232

Qy 752 LeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPhe 771

Db 2233 CTTGGTGTGTATTTCGCCATTTACGCAGTTGTTTGTTCATCTCGTGATATTTGTGTTT 2292

Qy 772 LeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaVal 791

Db 2293 TTGAAGTTCCAGAAACAAAGGCGATGCCCTTTGAAGTCATCTCTGAATTCCTTTCTGTT 2352

Qy 792 GlyAlaLysGlnAlaAla 798

Db 2353 GGAGCAAGCAGGCTGCTTCT 2373

RESULT 10

US-09-938-842A-1315

; Sequence 1315, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; TITLE OF INVENTION: SAME, AND METHODS OF USE

; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 1315

; LENGTH: 2190

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana									
US-09-938-842A-1315									
Alignment Scores:									
pred. No.:		8.48e-246		Length:		2190			
Score:		2584.00		Matches:		512			
Percent Similarity:		80.78%		Conservative:		89			
Best Local Similarity:		68.82%		Mismatches:		123			
Query Match:		63.41%		Indels:		20			
DB:		9		Gaps:		8			
US-10-051-909-32 (1-800) x US-09-938-842A-1315 (1-2190)									
Qy	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75						
Db	1	ATGAGTGGAGCTGTGCTTGTGCTATTGCTGCTGCTGGCAACTTGTATACAGGATGG	60						
Qy	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn	95						
Db	61	GATAACGCAACTATTGCAGAGCTGTGTTGTACATAAAAAAGGAGTTTTAATTGGAGAGT	120						
Qy	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115						
Db	121	AATCCATCAGTGAAGGCTCAATTGTGGCGATGTCACTTATTGGTGCTACTCTGATTACA	180						
Qy	116	ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer	135						
Db	181	ACATGCTCTGGAGGGGTAGCTGATTGGCTTGGTCGCGTCCCATGCTAATAATTGTCTCA	240						
Qy	136	IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu	155						
Db	241	ATTCTCTACTTGTGTTGTTCTTAGTAATGCTATGCTCTCGAATGTTTATGTTGTGCTC	300						
Qy	156	LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr	175						
Db	301	TTAGGAAGGTTGTTAGATGGATTGGGGTTGGTCTGTGGTTCACACTTGTTCCTATTAT	360						
Qy	176	IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer	195						
Db	361	ATATCTGAGACTGCACCACTGAGATTAGGGGACTTGAATACGCTACCGCAGTTCACT	420						
Qy	196	GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer	215						
Db	421	GGCTCTGAGGGATGTTCTTATCTTACTGTATGGTTTTTCGGAATGTCGTTGATGCCATCA	480						
Qy	216	ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu	235						
Db	481	CCTAGCTGGAGATTGATGCTTGGTGTCCTTTTCATCCCTTCCCTTGTCTTTTTCTTCCTC	540						
Qy	236	ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu	255						
Db	541	ACGGTCTTCTTCTTGGCCCGAGTCCCCAAGGTGGCTCGTGAGCAAAGGTCGAATGCTTGA	600						
Qy	256	AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu	275						
Db	601	GCAAAGCGGGTTCTTCAGAGACTGCGTGGTCGCCGAAGATGTGCTGGTGAGATGGCTTTG	660						
Qy	276	LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro	295						
Db	661	TTGGTTGAGGGTCTTGAATTGGAGGTGAACAACCATAGAGGAATATATAATTGGTCCC	720						
Qy	296	AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr	315						
Db	721	GCGGATGAAGTTACTGATGATCATGATATAGCTGTGGATAAGGATCAAATTAAAGTTATAT	780						
Qy	316	GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly	335						
Db	781	GGTGCAGAAAGGGCTGAGTTGGGTGCTAGGCCAGTCAAAGGA-----GGA	828						
Qy	336	SerValLeuSerLeuAlaSerArgHisGlySer---MetValAsnGlnSerValProLeu	354						
Db	829	AGCACTATGAGTGTTTGTCTCGCCATGGAAGTACAATGAGCAGGAGGCAAGGCTCATTTG	888						

QY	355	MetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGly	374						
Db	889	ATTGATCCTCTTGTCCACACTGTTTGGAGCGTTTCCGAGAAGATCCCGACACT--GGA	945						
QY	375	SerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHis	394						
Db	946	AGCATGAGGAGTGCCTTGTTCACATTTTGGAGTATGTTCAAGTGTGGAGGGAATCAA	1005						
QY	395	AlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSer	414						
Db	1006	CCAAGACATGAAGATTGGGATGAAGAGAATCTTGTGGAGAAGGTGAGGATTATCCATCC	1065						
QY	415	AspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla	434						
Db	1066	GAC--CATGGAGATGATTCTGAAGATGATCTTCAATCTCCGTGATCTCACGTCAAAACG	1122						
QY	435	ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet	454						
Db	1123	ACAAGCATGGAG--AAAGACATGCCTCACACTGCTCATGGAACTCTTTCTACCTTCAGA	1179						
QY	455	ArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly	474						
Db	1180	CATGGAAGTCAAGTGCAGGAGCTCAAGGGGAAGAGCGGGTAGTATGGGATTGGAGGT	1239						
QY	475	GlyTyrGlnLeuAlaTyrLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly	494						
Db	1240	GGATGGCAAGTGGCATGGAATGGACGGAAGAGAGATGAATCGGGACAGAAAGAA--	1296						
QY	495	GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle	514						
Db	1297	-----GAAGGTTTCCCAGGATCTCGACGTGGCTCAATT	1329						
QY	515	ValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAla	534						
Db	1330	GTTTCATTGCCTGGTGGTATGGAACCGGTGAG--GCAGATTGTGTACAAGCGTCTGCT	1386						
QY	535	LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla	554						
Db	1387	TTGGTTAGCCAAACAGCTCTTTATTCCAAAGACCTTCTCAAAGAACATACAATTGGTCCT	1446						
QY	555	AlaMetValHisProSerGluValAlaAlaLysGlySerArgTyrLysAspLeuPheGlu	574						
Db	1447	GCTATGGTACATCCATCCGAA---ACAACTAAAGGGTCAATTGGCATGATCTTTCATGAT	1503						
QY	575	ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAla	594						
Db	1504	CCTGGAGTCAAGCGTGCATTAGTCGTAGGAGTTGGACTTCAATACTTTCAGCAGTTCTCA	1563						
QY	595	GlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal	614						
Db	1564	GGCATCAACGGAGTCTTTACTACACACCCGAAAATCCTTGAGCAGCGGGGTGTCGGGATC	1623						
QY	615	IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr	634						
Db	1624	CTACTATCGAACATGGGGATTAGTTCTTCTCCTCAGCATCCTTACTTATAAGTGCATTGACA	1683						
QY	635	ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg	654						
Db	1684	ACCTTTGTGATGTTACTGCAATAGCTGTGCAATGAGGCTCATGGATCTTCTGGTTCGA	1743						
QY	655	ArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal	674						
Db	1744	AGGACCTTGCTTCTCACACAGATACCAATCCTGATAGCATCTCTATTTGGTTTTAGTAATC	1803						
QY	675	SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle	694						
Db	1804	TCAAAATCTTGTTCACATGAACAGCATTTGTGCACGGGTCTTATCAACCGTAAGCGTTGTG	1863						
QY	695	ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu	714						
Db	1864	CTCTACTTCTGCTTCTCGTGATGGGTTTCGGTCTCTGCTCCAAACATCTCTCTGTTCAGAG	1923						
QY	715	IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle	734						



[illegible]

RESULT 11  
US-09-938-842A-1315  
; Sequence 1315, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIP1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1315  
; LENGTH: 2190  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1315

Alignment Scores:	
Pred. No.:	8.48e-246
Score:	2584.00
Percent Similarity:	80.78%
Best Local Similarity:	68.82%
Query Match:	63.41%
DB:	11
	Length: 2190
	Matches: 512
	Conservative: 89
	Mismatches: 123
	Indels: 20
	Gaps: 8

US-10-051-909-32	(1-800)	x US-09-938-842A-1315	(1-2190)
Oy	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr	75
Dd	1	ATGAGTGGAGCTGTGCTTGTTGCTATTGCTGCTGCTGCTGGCAACTTGTTCACAAGCATGG	60
Oy	76	AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGlupheGlnLeuGlnAsn	95
Dd	61	GATAACGCCAACTATTGCAGGAGCTGTGTTGTACATAAAAAGGAGTTTTTAATTGGAGAGT	120
Oy	96	GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr	115
Dd	121	AATCCATCAGTGAAGGTCTAATTGTGGCGATGTCACCTTATTGGTGCTACTCTGTATTACA	180
Oy	116	ThrpheSerGlyProLeuSerAspSerIleGlyArgArgPrometLeuIleLeuSerSer	135
Dd	181	ACATGCTCTGGAGGGGTAGCTGATTGGCTTGGTCGCCGTCCCCTGCTAATAATTGTCTCTCA	240
Oy	136	IleLeuTyrrphepheserGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu	155

Db 1297 -----GAAGGTTTCCAGGATCTCGACGTGGCTCAATT 1329

QY 515 ValSerLeuProGlyGlyAspValPheGluGlySerGluPheValHisAlaAla 534  
|||||  
Db 1330 GTTTCATTGCCTGGTGGTATGGAACCGGTGAG--GCAGATTTGTACAAGCGTCTGCT 1386  
|||||

QY 535 LeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAla 554  
|||||  
Db 1387 TTGGTTAGCCAACCGACTCTTTATTCCAAAGACCTTCTCAAAGAACATACAATTGGTCCT 1446  
|||||

QY 555 AlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGlu 574  
|||||  
Db 1447 GCTATGGTACATCCATCCGAA--ACAACTAAGGGTCAATTTGGCATGATCTTCATGAT 1503  
|||||

QY 575 ProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnPheAla 594  
|||||  
Db 1504 CCTGGAGTCAAGCGTGCAATAGTCGTAGGAGTTGGACTTCAAATACTTCAGCAGTTCTCA 1563  
|||||

QY 595 GlyIleAsnGlyValLeuTyrThrProGlnIleLeuGluGlnAlaGlyValAlaVal 614  
|||||  
Db 1564 GGCATCAACGGAGTTCTTTACTACACACCGCAATCCTTGAGCAGCGGGTGTTCGGGATC 1623  
|||||

QY 615 IleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThr 634  
|||||  
Db 1624 CTTACTATCGAACATGGGGATTAGTTCTTCTCCTCAGCATCCTTACTTATAAGTGCATTGACA 1683  
|||||

QY 635 ThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArg 654  
|||||  
Db 1684 ACCTTTGTGATGTTACCTTGCAATAGCTGTTGCAATGAGGCTCATGGATCTTCTGTGTCGA 1743  
|||||

QY 655 ArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValVal 674  
|||||  
Db 1744 AGGACCTTGCTTCTCACACGATACCAATCCTGATAGCATCTCTATTGGTTTTAGTAATC 1803  
|||||

QY 675 SerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIle 694  
|||||  
Db 1804 TCAAACTTGTTCACATGAACAGCATTGTGCACCGGTCTTATCAACCGTAAGCGTTGTG 1863  
|||||

QY 695 ValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGlu 714  
|||||  
Db 1864 CTCTACTTCTGCTTCTTCGTGATGGGTTTTCGGTCTGCTCCAAACATCCTCTGTTTCAGAG 1923  
|||||

QY 715 IlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIle 734  
|||||  
Db 1924 ATTTTCCAACTCGAGTCCGCGGAATCTGCATCGCCTCTGCGCACTCACCTTCTGGATC 1983  
|||||

QY 735 GlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGly 754  
|||||  
Db 1984 TGTGACATAAATCGTCACTTACAGTCTCCCGTGCTGCTCAAATCCATTGGACTAGCTGGT 2043  
|||||

QY 755 ValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysVal 774  
|||||  
Db 2044 GTGTTTGGAAATGTACGCAATCGTATGTTGCATTTTCATGGGTCTTTGIGTTCATTAAAGTC 2103  
|||||

QY 775 ProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAlaLys 794  
|||||  
Db 2104 CCGGAACCTAAAGGCATGCCACTTGAAGTCATCACAGAGTTCTTTTCTGTTGGAGCTAGA 2163  
|||||

QY 795 GlnAlaAlaAla 798  
|||||  
Db 2164 CAAGCTGAAGCT 2175  
|||||

RESULT 12  
US-10-424-599-10324  
; Sequence 10324, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 10324  
; LENGTH: 2671  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_109333C.1  
US-10-424-599-10324

Alignment Scores: 2.26e-244 Length: 2671  
Pred. No.: 2571.00 Matches: 517  
Score: 80.73% Conservative: 99  
Percent Similarity: 67.76% Mismatches: 107  
Best Local Similarity: 63.09% Indels: 40  
Query Match: 17 Gaps: 12  
DB:

US-10-051-909-32 (1-800) x US-10-424-599-10324 (1-2671)

QY 55 LysMetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGly 74  
:::|  
Db 244 GAAATGAAAGTACCGTCTCTCGTCGCTATTGCCGTGCTATCGGCAATATCTCCAGGGA 303  
|  
QY 75 TrpAspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGln 94  
|  
Db 304 TGGGATAATGCTTCCATCGCCGGGGCCATTGTTTACATTAAAGAAAGACCTTGCCTGCAA 363  
|  
QY 95 AsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleVal 114  
|  
Db 364 ACA-----ACTATGGAAGGGCTTGTGTGGCCATGTCCCTGATTTGGAGCAACGGTAATC 417  
|  
QY 115 ThrThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSer 134  
|  
Db 418 ACCACATGCTCTGGTCTTATAGCGGATTTGGCTCGGTCCGGACCCCATGATGATAATCTCA 477  
|  
QY 135 SerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeu 154  
|  
Db 478 TCTGTGCTCTATTCTTTGGGTGGTTTGGTGATGCTGTGGTCCCAAATGTGTATGTGTG 537  
|  
QY 155 LeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeu 174  
|  
Db 538 TGCTTTGGCGAGGCTACTTGTATGGATTTGGGATTTGGCTTGTGTGACTCTTTGTCCCGGTC 597  
|  
QY 175 TyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPhe 194  
|  
Db 598 TATATATCTGAAACGGCGCGCTCTGAAATAAGGGGGTCTGTTGAATACGCTTCTCCTCAGGTC 657  
|  
QY 195 SerGlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerPro 214  
|  
Db 658 AGTGGCTCTGGAGGAATGTTTGTGCTGCTTATGTTGTTTGGCATGTTCATTGAGTCCC 717  
|  
QY 215 SerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGly 234  
|  
Db 718 GCGCCTAGCTGGAGGCTCATGCTTGGGTTCTGTCTATTCCCTCCCTCTTGTATTTTGCA 777  
|  
QY 235 LeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMet--- 253  
|  
Db 778 TTGACCATTTTTTTTCTTGCCCGAGTCTTCC-----TCTTCAGGCTCGGTCCAGC 825  
|  
QY 254 -----AlaGluAlaLysLysValLeu-GlnLysLeuArgGlyLysAs 267  
|  
Db 826 AAAATGTAATTTTGATGCTGCGAGCGGCAACAAAGTGTCCCAAAGATTGCGCGGACGGGA 885  
|  
QY 267 pAspValSerGlyGluLeuSerLeuLeuGluGly-LeuGlu-ValGlyGlyAspThr 286  
|  
Db 886 GGATGTGTCTGGCGGAGATGACCTTGTCTGGTTGAAGGATCTACGGGATTTGGGGTGATACA 945  
|  
QY 287 SerIleGluGluTyrIleIleGly-ProAlaThrGluAlaAlaAspLeuValThrAs 306  
|  
Db 946 TCTATTGAAGAGTACATAATTGGACCCCTGCTGACCCAGGTGGCTGATGGTCAATGAACATGC 1005  
|

Qy 306 pGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaAr 326  
Db 1006 AACAGAGAAAGATAAAATTCGATTATATGGATCCCAAGCAGCGCTCTCTGGTTAGCAAA 1065  
Qy 326 gProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySe 346  
Db 1066 ACCTGTCACTGGACAGAGTTCTATTGGC-----CTTGCAACGCCATGGAAG 1113  
Qy 346 rMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHi 366  
Db 1114 CATCATCAACCAAGCATGCCCTCATGGATCCTCTGGTGACACTGTTGGTAGCATTCA 1173  
Qy 366 sGluAsnMetProGln-----AlaGlyGlySerMetArgSerThrLeuPheProAsnPh 384  
Db 1174 TGAGAAGCTCCCGAGACAGGAGCAGGAGGAGCATGCGAAGCACTCTGTTTCCAAATTT 1233  
Qy 384 eGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAs 404  
Db 1234 TGGAAGCATGTTTCAGCACTGCTGAGCCGCATGCTAATAAATGAACAGTGGGATGAAGAG 1293  
Qy 404 nLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAs 424  
Db 1294 CTTACAAAGGGAAGGTGAGGACTACATGTCAGATGCAGCCGTGGGACTCTGATGATAA 1353  
Qy 424 nLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle---- 442  
Db 1354 TTTGCAAGTCCTTTAATCTCACGCCCAACAACAGCCTTGAA---AAAGACTTGCCTCC 1410  
Qy 443 -----ValHisHisGlyHisArgGlySerAlaLeu---SerMetArgArgGlnThrLe 459  
Db 1411 TCCTCCTCCTTCCCAT-----GGCAGTATCCTTGGCAGCATGAGGGGTCACAGTAG 1461  
Qy 459 uLeuGlyGluGly---GlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLe 478  
Db 1462 TCTCATGCAAGGTCAGGTGAGCAAGGTGGTAGTACAGGTATTGGTGGGTGGCAACT 1521  
Qy 478 uAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysAr 498  
Db 1522 GGCATGGAAATGGACTGATAA---GGTGAGGATGGAAACAACAAGAGGGGTTTAAAG 1578  
Qy 498 gValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuPr 518  
Db 1579 GATTTATTATACATGAGGAGGAGTTTCTGTCATCTCGTCGTGGATCCATTGTATCGATTCC 1638  
Qy 518 oGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaAlaLeuValSerGl 538  
Db 1639 CGGTGAAGGC-----GAATTGTCCAGGCTGCTGCCTTGGTAAAGCCA 1680  
Qy 538 nSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHi 558  
Db 1681 ACCCGCTCTTTACTCCAAGGAGCTTATTGATGGACACCCAGTTGGGCGCTGCAATGGTTCA 1740  
Qy 558 sProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValAr 578  
Db 1741 CCCATCTGACAGAGCTTCAAAGGGGCCAAGTTGGAAAGCTCTTCTTGAACCAAGGGTTAA 1800  
Qy 578 gArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGl 598  
Db 1801 GCATGCATTGGTTGTTGGAGTTGGAATACAATACTTCAGCAGTTCAGGGGATAAATGG 1860  
Qy 598 yValLeuTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLy 618  
Db 1861 GGTTCATATTACACACCTCAAATCCTTGAAGAGGCCGGTGTGTGAAGTCTCTTCTTCAGA 1920  
Qy 618 sPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMe 638  
Db 1921 TATAGGCATTGGCTCAGAGTCGGCATCATTCCTTATCAGTGCTTTTCACAACCTTCTTGAT 1980  
Qy 638 tLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyVArgArgPheLeuLe 658  
Db 1981 GCTTCCTGTATAGCGGTAGCCATGAAGCTCATGGATGTTTTCAGGCAGAAGGCAGTTGTT 2040

Qy 658 uLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuValSerAsnLeuIl 678  
Db 2041 ACTTACTACAATCCCGTGTCTGATGGGTCACTCATTTATTTGGTCAATGGAGCCTGGT 2100  
Qy 678 eAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCy 698  
Db 2101 AAATTTTGGCAATGTGCCCCATGCAGCAATCTCAACAGTATGCGTTGTGTTTATTCTG 2160  
Qy 698 sCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProTh 718  
Db 2161 CTGCTTTGTGATGGGTTATGGACCAATTCCAAACATCCTTTGCTCAGAGATTTTCCCCAC 2220  
Qy 718 rArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIl 738  
Db 2221 TAGGGTGGGTGGCCTCTGCAATTGCTATCTGTGCATTAGTGTCTCGATTGGAGACATCAT 2280  
Qy 738 eValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIl 758  
Db 2281 CATCACATACTCGCTGCCTGTGATGCTGGGCTCTTTAGGACTTGGTGTGTATTGCGCAT 2340  
Qy 758 eTyrAlaValValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLy 778  
Db 2341 TTACGCAGTTGTTGTTTCATCTCGTGGATATTGTGTGTTTTGAAGGTTCCAGAAACAA 2400  
Qy 778 sGlyMetProLeuGluValIleThrGluPhePheAlaValAlaGlyAlaLysGlnAlaAl 798  
Db 2401 GGGCATGCCCTTGAAGTCATCTCTGAATTCTTTTCTGTGGAGCAAGCAGGCTGCTTC 2460  
Qy 798 a 798  
Db 2461 T 2461  
RESULT 13  
US-10-425-114-7958  
; Sequence 7958, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 7958  
; LENGTH: 2405  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700728773\_FLI  
US-10-425-114-7958  
Alignment Scores:  
Pred. No.: 1.16e-241 Length: 2405  
Score: 2543.00 Matches: 508  
Percent Similarity: 79.54% Conservative: 83  
Best Local Similarity: 68.37% Mismatches: 120  
Query Match: 62.40% Indels: 32  
DB: 17 Gaps: 10  
US-10-051-909-32 (1-800) x US-10-425-114-7958 (1-2405)  
Qy 56 MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp 75  
Db 76 ATGAGCGGAGCTGTTCTTGTGTGCTGGCTGCTGCTATCGGCAACTTGCTGCAAGGATGG 135  
Qy 76 AspAsnAlaThrIleAlaAlaValLeuTyrIleLysLysGluPheGlnLeuGlnAsn 95



Dbb 136 GATAATGCCACCATTCGAGGATCTATTTTGTACATAAAGAGGGAGTTTAATTTACAAAGT 195  
QY 96 GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr 115  
Dbb 196 GAACCCACAATAGAAGGTCTAATTTGGCCATGTCACCTATTATGGAGCCACTGTCGTTACG 255  
QY 116 ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer 135  
Dbb 256 ACATGCTCTGGACCCCTTATCAGACTTGTAGCCGCGCTCCTATGTTGATCATCTCCTCT 315  
QY 136 IleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnValTyrValLeuLeu 155  
Dbb 316 ATCCTTTATTTTGTGGTTCCTGCTCATGCTATGGTCTCCAAATGTTTATATCTCCTC 375  
QY 156 LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr 175  
Dbb 376 TTTGCAAGGCTTTTAGATGGATTAGGCATTGGTTTGGCGGTACCTTGGTACCTCTTTAC 435  
QY 176 IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer 195  
Dbb 436 ATATCCGAGACAGCTCCACCTGAGATTAGGGGATTACTAAATACTCTTCGCGAGTTCACT 495  
QY 196 GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer 215  
Dbb 496 GGTTCTGCTGGAATGTTCTTTTCTCTACTGTATGGTCTTTGCTATATCACTCACCAAGGCC 555  
QY 216 ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu 235  
Dbb 556 CCAAACTGGAGACTCATGTTAGGTGTTCTTTTCAATCCCTCTCTCATTTATTTTGCACTC 615  
QY 236 ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu 255  
Dbb 616 ACACATATCTTCTTGCTGCTGAATCTCCAAGATGGCTGTCTCAGCAAAGCCGCGATGTAGAG 675  
QY 256 AlaLysLysValLeuGlnLysLeuArgGlyLysAspValSerGlyGluLeuSerLeu 275  
Dbb 676 GCCAAGAAGGTTTGTCAACGACTTCGCGGCAGACAAAGATGTCGTGGTGAGATGGCTTTA 735  
QY 276 LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro 295  
Dbb 736 CTAGTTGAGGGTCTTGGAGTTGGGGTGATACCGTTATAGAGAGTACATAATAGTCCA 795  
QY 296 AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr 315  
Dbb 796 GCCAATGAATTCAGT-----GATGCAGAGCAGATCAAAATTTGTAT 834  
QY 316 GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly 335  
Dbb 835 GGGACAGCAGAAGGTCAATCCTGGATTGCCAAACCTGTCACTGGACAAAGTCCATTGGC 894  
QY 336 SerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMet 355  
Dbb 895 -----CTTGATCTCGGAAGGGAAGCATGGCAAATCAAAGTGCT--CTAGTG 939  
QY 356 AspProIleValThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySer 375  
Dbb 940 GACCCCTCTAGTGAAGCTCTTTGGTAGTGTCATGAGAAGCTCCAGAAACAGGA----- 993  
QY 376 MetArgSerThrLeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAla 395  
Dbb 994 -----AGCACACTTTCCACACTTTGGAAGCATGTTTAGTGTGGGGGAAATCAACCT 1047  
QY 396 LysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAsp 415  
Dbb 1048 AGGAATGAAGATTGGGATGAAGAAAGCATTCGACAGAGGGAGATGATTATGTATCTGTAT 1107  
QY 416 GlyAlaGly---GlyAspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAla 434  
Dbb 1108 GCTGCTGACACCGACGATTCTGTATGACAATTTGCAGAGTCCATTGATCTCACGTCAAGCA 1167  
QY 435 ThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMet 454  
Dbb 1168 ACAAGTGGGAA---AGGGACATGCCTGCTCTGTCCCGAGGA-----AGCATG 1212

QY 455 ArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGly 474  
Dbb 1213 AGGCAAGGTACTCTTTTG-----CAAGGAGAACCTGCTCGAAAACTCAGGGATTGGTGGT 1266  
QY 475 GlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGly 494  
Dbb 1267 GGATGGCAGCTAGCATGGAAATGGTCTGAAACAGAAAGGTGT----- 1308  
QY 495 GlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIle 514  
Dbb 1309 ---TTCAAGAGAATATATTTTGCACCAAGAAGGTGGTCTCTGGATCTAGACGTGGGTCCATC 1365  
QY 515 ValSerLeuProGly--GlyGlyAspValPheGluGlySerGluPheValHisAlaAla 533  
Dbb 1366 ATTTCACTTCCAGGCTGTGATGCGCGACACTAAACAGATGGTGAGATTGTCCAGGCTGCC 1425  
QY 534 AlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAsp 553  
Dbb 1426 GCTCTAGTGAGTCAATCAGCCCTTTTACAACAAGGAGCTTATGCATCAGCAGCCAGTTGGA 1485  
QY 554 AlaAlaMetValHisProSerGluValAlaAlaLysGlySerArgTrpLysAspLeuphe 573  
Dbb 1486 CCAGCCATGATTCAATCCATCCCAAACAGCTGCAAAAGGCCGAGTTGGAGTGACCTTTTA 1545  
QY 574 GluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIleLeuGlnGlnPhe 593  
Dbb 1546 GAGCCTGGAGTGAAGCATGCACCTGATTGTGGGGTGGGGATTCAAAATCTTCAACAGTTC 1605  
QY 594 AlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAla 613  
Dbb 1606 TCTGGTATAAATGGGGTGTCTCTACTATACACCTCAAATTTCTTGAGCAGCAGGCTGTTGGT 1665  
QY 614 ValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeu 633  
Dbb 1666 TATCTTCTTTCAAACCTGGGCCTCGGTTCTACTTCTGCGTCTTTTCTTATCAGCAGCGTG 1725  
QY 634 ThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGly 653  
Dbb 1726 ACAACCTTGTGTGATGTTACCTGTATAGCTGTGGCCATGAGGCTCATGGATATATCTGGC 1785  
QY 654 ArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIleLeuVal 673  
Dbb 1786 AGAAGACCTTGTGCTGACTACAATCCCGTCTTGATAGTATCATCTTCATATTAGTA 1845  
QY 674 ValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerVal 693  
Dbb 1846 ATAGGGAGTCTTGTGGAAATTGGACAGTACTATAAATGCTTTTATCTCAACTTCGAGTGT 1905  
QY 694 IleValTyrPheCysPheValMetGlyPheGlyProIleProAsnIleLeuCysAla 713  
Dbb 1906 ATTGCTACTTCTGTGTTTCGTCAATGGGCTTTGGGCCAAATTCCTAATATCCTTTGTTCA 1965  
QY 714 GluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThrPheTrp 733  
Dbb 1966 GAGATCTTCCCTACTCGAGTTCGTGGTCTATGCATAGCTATATGIGCTCTTACATTTTGG 2025  
QY 734 IleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGlyLeuAla 753  
Dbb 2026 ATATGTGACATCATAGTCACCTACTCGCTCCAGTTATGCTTAAITCTGTGGGCCTTGGT 2085  
QY 754 GlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPheValPheLeuLys 773  
Dbb 2086 GGTGCTCTTGGTATGTATGCGGTTGTGTGCATCATAGCTATGTTGGTGTGTTCTTCTTAAAA 2145  
QY 774 ValProGluThrLysGlyMetProLeuGluValIleThrGluPhePheAlaValGlyAla 793  
Dbb 2146 GTTCAGAAACCAAGGCGATGCCACTGGAAAGTTATCATTTGAGTTCITCTCTGTCTGGGGCA 2205  
QY 794 LysGlnAla 796  
Dbb 2206 AAGCAAGCT 2214

RESULT 14  
US-10-424-599-108435  
; Sequence 108435, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 108435  
; LENGTH: 2573  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_68931C.1  
US-10-424-599-108435

Alignment Scores:				
Pred. No.:	1,04e-235	Length:	2573	
Score:	2483.50	Matches:	505	
Percent Similarity:	76.69%	Conservative:	84	
Best Local Similarity:	65.76%	Mismatches:	121	
Query Match:	60.94%	Indels:	58	
DB:	17	Gaps:	12	
US-10-051-909-32 (1-800) x US-10-424-599-108435 (1-2573)				
Qy	56	MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTrp	75	
Db	37	ATGAGCGGAGCTGTTCTTGTGTGCTGCTGCTATCGCAACTTGTGCAAGGATGG	96	
Qy	76	AspAsnAlaThrIle-----	80	
Db	97	GATAATGCCACCATTGCAGGTCAATATTTTATTCGGTTAGCCAATTAAGATAAATTCAA	156	
Qy	81	-----AlaAlaAlaValLeuTyrIleLysLysGluPhe	91	
Db	157	CCAAGTCTAATTGSGCTTCTATGCTATGCAGGATCTATTTTGTACATAAAGAGGGAGTTT	216	
Qy	92	GlnLeuGlnAsnGluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAla	111	
Db	217	AATTTACAAAGTGAACCCACAATAGAGGTCTAATTGTGGCCATGTCACCTATTGGAGCC	276	
Qy	112	ThrIleValThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeu	131	
Db	277	ACTGTGTTACGACATGCTCTGGACCCCTTATCAGACTTGCTAGGCCGGCGTCCTATGTTG	336	
Qy	132	IleLeuSerSerIleLeuTyrPhePheSerGlyLeuIleMetLeuTrpSerProAsnVal	151	
Db	337	ATCATCTCCTCTATCCTTTATTTGTGGTTCTCTGGTCACTGCTATGGTCTCCAAATGTT	396	
Qy	152	TyrValLeuLeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeu	171	
Db	397	TATATTCTCCTCTTTGCAAGGCTTTTAGATGGATTAGGCATTGGTTTGGCGGTACCTTG	456	
Qy	172	ValProLeuTyrIleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeu	191	
Db	457	GTACCTCTTTACATATCCGAGACAGCTCCACCCTGAGATTAGGGGATTACTAAATACTCTT	516	
Qy	192	ProGlnPheSerGlySerGlyMetPheLeuSerTyrCysMetValPheGlyMetSer	211	
Db	517	CCGAGTTCACTGGTCTCTGCTGGAATGTTCTTTCTCTACTGTATGGTCTTTGCTATGTCA	576	
Qy	212	LeuSerProSerProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhe	231	
Db	577	CTCACCAAGGCCCCCAAATTTGGAGACTCATGTTGGGTGTTCTTTTCAATTCCTCTCTCATT	636	

Qy	232	PhePheGlyLeuThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGly	251	
Db	637	TATTTTGCACTCAGCTATTCTTCTTGCTGAATCTCCAAGATGGCTTGTC--AAAGGC	693	
Qy	252	ArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGly	271	
Db	694	CGGATGCTAGAGGCCAAGAAGTTTTTGCACGACTTCGCGGCAGACAGAATGTCGCTGT	753	
Qy	272	GluLeuSerLeuLeuGlyLeuGluValGlyGlyAspThrSerIleGluGluTyr	291	
Db	754	GAGATGGCTTACTAGTTGAGGGTCTTGAGTTGGCGTGATACCGCTATAGAAGATAC	813	
Qy	292	IleIleGlyProAla-ThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGl	311	
Db	814	ATAATTGTCAGCTCAATGAATTCAGT-----GAGGCAGAGCA	852	
Qy	311	nIleThrLeuTyrGlyProGluGlyGlnSerTrpIleAlaArgProSerLysGlyPr	331	
Db	853	GATCAAATTTGATGGGACAGCAGAAAGGTGTATCTCTGGATTGCTAAACCTGTCACTGGACA	912	
Qy	331	oIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsnGlnSe	351	
Db	913	AAGTTCCATGGT-----CTTGTATCTCGGAAGGGAAGCATGGCAAATCAAAG	960	
Qy	351	rValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMetProGl	371	
Db	961	T---GGTCTAGTGGACCCTCTAGTGAACCTCTTTGGTAGTGTCCATGAGAAGCTCCCAGA	1017	
Qy	371	nAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSerValTh	391	
Db	1018	AACAGGA-----AGCACACTTTTCCACACTTTTGGAAAGCATGTTTAGTGTGG	1065	
Qy	391	rAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAspGluGl	411	
Db	1066	GGGGAATCAACCTAGGAATGAAGATTGGGATGAAGAAAGCATTTGCCAGAGAGGGAGATGA	1125	
Qy	411	uTyrAlaSerAspGlyAlaGly---GlyAspTyrGluAspAsnLeuHisSerProLeuLe	430	
Db	1126	TTATGTATCTGATGCTGTTGACACCCGACGATTCTGATGACAAATTTGAGAGTCCATTGAT	1185	
Qy	430	uSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGlySe	450	
Db	1186	CTCACGTCAAGCAACAAGTGGGAA---AGGGACATGCCTGCTCTGCCAGGGA-----	1237	
Qy	450	rAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAspGlyValSerSerTh	470	
Db	1238	-----AGCATGAGGCNAAGGTAGTCTTTTG-----CAAGGAGAACCTGCTGGAAACTC	1284	
Qy	470	rAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGl	490	
Db	1285	AGGGATTGGTGGTGGATGGCAGCTAGCATGGAAATGGTCTGMAACAGAAAGGTGTT-----	1339	
Qy	490	yArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySerAr	510	
Db	1340	-----TTCAAGAGAATATATTGCAACCAAGAGGTGGTCTCTGGATCTAG	1383	
Qy	510	gArgGlySerIleValSerLeuProGly---GlyGlyAspValPheGluGlySerGluPh	529	
Db	1384	ACGTGGGTCCATCATTTTCACCTCCAGGCTGTGATGCCCGCACACTAACAGATGGTGAGAT	1443	
Qy	529	eValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeuAlaGluPr	549	
Db	1444	TGTCCAGGCTGCCGCTCTAGTGAGTCAATCAGCCCTTTTACAACAAGAGGCTTATGCATCA	1503	
Qy	549	oArgMetSerAspAlaAlaMetValHisProSerGluValAlaLalysGlySerArgTr	569	
Db	1504	GCAGCCAGTTGGACCAGCCATGATTTCATCCATCCCAAACAGCTGCAAAAGGGCCGAGTTG	1563	
Qy	569	pLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGlyIleGlnIl	589	
Db	1564	GAGTGACCTTTTAGAGCCTGGAGTGAAGCATGCATGATTGTGGGGGTGGGGATTCAAAT	1623	
Qy	589	eLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGl	609	

```

Db      1624 TCTTCAACAGTTCTCTGGTATAAATGGGGTCTCTACTATACACCTCAAATCTTTGAGCA 1683
QY      609 nAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLe 629
Db      1684 GGCAGGTGTTGGTTATCTTCTTCAAACCTGGGCTCGGTCTCTACTTCTGCGTCTTTTCT 1743
QY      629 uIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMe 649
Db      1744 TATCAGCAGCGTGACAACCTTGTTGATGTTACCCTGTATAGTGTGGCCATGAGGCTCAT 1803
QY      649 tAspLeuSerGlyArgArgPheLeuLeuGlyThrIleProIleLeuIleAlaSerLe 669
Db      1804 GGATATATCTGGCAGAGAACCTTGCTGCTGACTACAATCCCCGTCTTGATAGTATCACT 1863
QY      669 uValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSe 689
Db      1864 TCTCATATAGTAATAAGGAGCTTGTGGAATTGGACAGTACTATAAATGCTTTTATCTC 1923
QY      689 rThrValSerValIleValIleValIleValIleValIleValIleValIleValIleVal 709
Db      1924 AACTTCGAGTGTATTGTCTACTTCTGTTGTTTCGTCTCATGGGCTTTGGGCCAATTCCTAA 1983
QY      709 nIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAl 729
Db      1984 TATCCTTTGTTTCAGAGATCTTCCCTACTCGAGTTCGIGTCTATGCATAGCTATATGTGC 2043
QY      729 aPheThrPheThrIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAl 749
Db      2044 TCTTACATTTTGGATATGTGACATCATAGTCACCTACTCGCTCCAGTTATGCTTAATTC 2103
QY      749 aIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSerPheValPh 769
Db      2104 TGTGGCCCTTGGTGGTGTCTTTGGTATGTATGCGGTTGIGTGCATCATAGCTTGGGTGTT 2163
QY      769 eValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThrGluPhePh 789
Db      2164 TGTCTTCTTAAAGTTCAGAAACCAGGGCATGCCACTGGAAGTTATCATTTGAGTTCTT 2223
QY      789 eAlaValGlyAlaLysGlnAla 796
Db      2224 CTCTGTGGGGCAAAGCAAGCT 2245

RESULT 15
US-10-425-114-29146
; Sequence 29146, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 29146
; LENGTH: 2271
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB83-005-C10_FLI
US-10-425-114-29146
```

```

Alignment Scores:      1.01e-231      Length:      2271
Pred. No.:             2442.50        Matches:      472
Score:                 83.66%          Conservative:  76
Percent Similarity:    72.06%          Mismatches:   100
Best Local Similarity:
```

```

Query Match:      59.94%      Indels:      7
DB:               17         Gaps:       6

US-10-051-909-32 (1-800) x US-10-425-114-29146 (1-2271)

QY      143 LeuIleMetLeuTrpSerProAsnValTyrValLeuLeuLeuAlaAargPheValAspGly 162
Db      3 CTGGTGATGCTCTGGGCGCCAAGTGTGTACATCTTGCTCCCTCGCAAGGCTCATTTGATGG 62
QY      163 PheGlyIleGlyLeuAlaValThrLeuValProLeuTyrIleSerGluIleAlaProSer 182
Db      63 TTCGGTATCGGTTTGGCGGTACACACTTGTTCCTCTCTACATCTCCGAGACTGCGCCGACA 122
QY      183 GluIleArgGlyLeuLeuAsnThrLeuProGlnPheSerGlySerGlyGlyMetPheLeu 202
Db      123 GACATTTCGTGGGTGTTTGAAACACGTTGCCAGATTCCCGAGTTCAGTGGGTGAGGGGATGTTCTC 182
QY      203 SerTyrCysMetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeu 222
Db      183 TCCTACTGCATGGTGTGGGATGTCCCTCATGCCAAACCTGATTGGAGGCTCATGTCTT 242
QY      223 GlyValLeuAlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGlu 242
Db      243 GGAGTTCTGTCGATCCCGTCACCTATTATTACTTTGGACTGACTGTCTTCTACTTGCCTGAA 302
QY      243 SerProArgTrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLys 262
Db      303 TCACCAAGGTGGCTTGTGAGCAAAGGAAGGATGGCTGAGCGCAAGAGAGAGTGTTCAAAGG 362
QY      263 LeuArgGlyLysAspAspValSerGlyGluLeuSerLeuLeuGlyLeuGluVal 282
Db      363 CTGCGGGGAAGAGAGATGCTCTCAGGGGAGATGGCTCTTCTAGTTGAAGGTTTGGGGGTC 422
QY      283 GlyGlyAspThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAsp 302
Db      423 GGTAAAGATACACGTTATTGAAGAATACATAATTGTCCTCCGATGATGAACCTTGCTGATGAA 482
QY      303 LeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGlyGlnSer 322
Db      483 GGGCTGGCTCCA--GATCCAGAGAAGATCAAACTATATATGACCTGAAGAGGCCCTATCT 539
QY      323 TrpIleAlaAargProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer 342
Db      540 TGGGTTGCCCGACCTGTTCCGGGACAAAGTGCTCTTGGAAAGCGGTTAGGTCTCATCTCT 599
QY      343 ArgHisGlySerMet---ValAsnGlnSerValProLeuMetAspProIleValThrLeu 361
Db      600 CGTCATGGGAGTATGGCGGCTAGTCAGGGTAAGCCCTCGTGGATCCTATGTCCTACTCTT 659
QY      362 PheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhe 381
Db      660 TTCGGAAGTGTTCATGAAAAGATGCCTGAGATCATGCGGAGCATGAGGAGCACATGTGTT 719
QY      382 ProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAsp 401
Db      720 CCCAACTTTGGCAGCATGTTTAGTTGTTGCCGACCCAGCAGCGGTGAAAGCTGACTGGGAC 779
QY      402 GluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyr 421
Db      780 GCCGAGAGT---CAAAGGGAAGGTGAAGATTATGCTTCGGATCATGGTGGCGATGACATT 836
QY      422 GluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAsp 441
Db      837 GAGGATAACCTCCAAAGCCCACCTTATTTCGTTCAGGCAACAAGTGTGGAAGGAAGGAG 896
QY      442 IleValHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGly 461
Db      897 ATCGCT-----GCACCTCATGGTAGCATATTGGGTGCTGTGGGAAGGAGCAGTAGCTTG 950
QY      462 GluGlyGlyAspGlyValSerSerThrAspIleGlyGlyThrGlnLeuAlaTrpLys 481
Db      951 CAGGGTGGGGAGGCAGTAAGCAGCATGGGCATTGGCGGAGGATGGCAGTTGGCGTGGAAA 1010
```



Qy	482	TrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeu	501
Db	1011	TGGACCGAGAGAGAGGGCGAAGATGGGCAAAAGGAAGGTGGCTTCCAGCGTATTACTTG	1070
Qy	502	HisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGly	521
Db	1071	CATGAGGAGGGCGTACAAGGCAAC--AGGGGTCTATATTGTCAATTACCA--GGCGGG	1124
Qy	522	AspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeu	541
Db	1125	GATGTTCTCTCGTGGTGGTTCATCCAGGCTGCAGCTCTTGTGAGCCAACCAAGCTCTT	1184
Qy	542	PheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGlu	561
Db	1185	TACTCTAAGGAACCTGCTGGAGCAACGTGCTGCTGGTCTCTGCGATGATGCATCTGAA	1244
Qy	562	ValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgAlaLeu	581
Db	1245	GCAGTTACTAAAGGTCCAAGATGGGCCGACCTATTTGAGCCTGGGTGAAGCATGCACTG	1304
Qy	582	LeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyr	601
Db	1305	TTTGTGGCATAGGAATACAGATCCTGTCAACAGTTTGTGGCATCAACGGCGTTCTCTAC	1364
Qy	602	TyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeu	621
Db	1365	TACACTCCTCAAAATTCTTGAGCAAGCAGCGTCGGTGTCTTCTGTGAAACCTCGGCCCT	1424
Qy	622	SerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCys	641
Db	1425	AACGCTTCTTCGGCATCAATCCTCATTAGCGCCCTGACGACCTTACTGATGCTCCCAAGC	1484
Qy	642	IleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuGlyThr	661
Db	1485	ATCGGCATTGCGATGAGGCTCATGGATATGTCCGGAAGGAGGTTTCTCCTCCTCGCGACG	1544
Qy	662	IleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGly	681
Db	1545	ATCCCAAGTCTAATAGTCGGCTACTCGCTCTGGTGGTGTCCAACATCGTGGACGTGGGG	1604
Qy	682	ThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheVal	701
Db	1605	GACGTGGCGACGGCGGCTCTCCACGGCCAGCGTCATAGTCTACTTCTGCTTCTTCGTC	1664
Qy	702	MetGlyPheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArg	721
Db	1665	ATGGGGTTCGGGCCCGTCCCCAACATCCTCTGCGCAGAGATCTTCCCCACCACGGTCCGC	1724
Qy	722	GlyLeuCysIleAlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyr	741
Db	1725	GGTGTCTGCATCGCCATCTCGCCCTGGGCTTCTGGGTCCGGCAGACATCATCGTGACGTAC	1784
Qy	742	SerLeuProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaVal	761
Db	1785	ACTCTCCCGTGATGCTGAACGTCTCGGGCTCGCCGGCTCTTTGGGGTGACGCCGTC	1844
Qy	762	ValCysLeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetPro	781
Db	1845	GTGTGCGTCTAGCCCTCGCGTTCGTGTTTCATCAAGGTGCCCCGAGACGAAGGCATGCCT	1904
Qy	782	LeuGluValIleThrGluPhePheAlaValGlyAlaLysGlnAla	796
Db	1905	CTCGAGGTCAATCACCGAGTTCTTCTCCGTGGGGCAAGCAAGCC	1949

This page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: April 13, 2005, 12:44:19 ; Search time 5523.13 Seconds  
(without alignments)  
5513.431 Million cell updates/sec

Title: US-10-051-909-32  
Perfect score: 4075  
Sequence: 1 IRSGSWLAVQTPTFDLDR.....PLEVITEFAVGAKQAAKA 800

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q-/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074037\_14060/app\_query.fasta\_1.1678  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909@cgn 1 1 6628 @runat\_13042005\_074037\_14060 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_htc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3027	74.3	2180	3 AY105508	AY105508 Zea mays
2	2364	58.0	2292	9 CL980871	CL980871 OsIFCC045
3	1934.5	47.5	1932	9 CL968379	CL968379 OsIFCC040
C 4	1399	34.3	911	8 CC423708	CC423708 PUHOR49TB
5	1382.5	33.9	845	8 BZ723844	BZ723844 PUCFI60TD
C 6	1279	31.4	795	8 BZ989776	BZ989776 PUGDY07TD
7	1268.5	31.1	1135	3 CNS0A3H4	BX828912 Arabidops
8	1168	28.7	833	6 CA065405	CA065405 SCACAD103
9	1154	28.3	733	6 CA146057	CA146057 SCVPRT207

10	1142	28.0	696	6 CA075335	CA075335 SCJFAM106
11	1120	27.5	1082	7 CK208781	CK208781 FGAS02050
12	1101	27.0	829	5 BQ752737	BQ752737 WHE4118 E
13	1100	27.0	871	6 CA170109	CA170109 SCQGSB108
14	1086	26.7	1107	5 BX841312	BX841312 BX841312
15	1066.5	26.2	848	7 CF445664	CF445664 EST682009
16	1063.5	26.1	769	6 CA243478	CA243478 SCQGFU407
17	1055	25.9	994	6 CA067439	CA067439 SCQGAD106
18	1034	25.4	636	6 CF018228	CF018228 QBM9a03.x
19	1033.5	25.4	950	7 CK263963	CK263963 EST710041
20	1030	25.3	692	6 CA112602	CA112602 SCQQLB106
21	1022	25.1	631	6 CA252785	CA252785 SCBFFL114
22	1018	25.0	787	7 CF182124	CF182124 JG-2 Chin
23	1014.5	24.9	643	6 CA263039	CA263039 SCPRLB202
24	1014.5	24.9	858	6 CA207025	CA207025 SCEQSB1C0
25	1009	24.8	644	6 CA279697	CA279697 SCCCFU800
26	1003	24.6	616	6 CA081000	CA081000 SCCCFAM200
27	997	24.5	606	6 CA230633	CA230633 SCJFFFL3C0
28	990.5	24.3	745	6 CB635485	CB635485 OSIIEB150
C 29	981	24.1	594	1 AI861088	AI861088 603011H02
30	980	24.0	620	6 CA167255	CA167255 SCCCSB100
31	966	23.7	669	6 CA073177	CA073177 SCEPAM105
32	957	23.5	817	8 BZ819234	BZ819234 PUGDY07TB
33	945	23.2	903	7 CF211894	CF211894 CGF100066
34	941.5	23.1	655	6 CA268224	CA268224 SCQSRF305
35	939	23.0	701	6 CD879843	CD879843 AZO4.1061
36	937	23.0	827	6 CA207188	CA207188 SCEQSB1C0
37	935	22.9	801	7 CK201790	CK201790 FGAS01031
38	930	22.8	859	4 BM817360	BM817360 HC106B05
39	928	22.8	749	5 BU040051	BU040051 PP LEa000
40	924	22.7	688	6 CA502649	CA502649 WHE4338 A
41	921.5	22.6	700	5 BQ870435	BQ870435 QGD9A19.Y
42	921.5	22.6	703	6 CA502490	CA502490 WHE4048 B
43	921	22.6	687	6 CD893252	CD893252 G118.123E
C 44	916	22.5	764	7 CO983244	CO983244 GM89018A2
45	916	22.5	847	6 CB981615	CB981615 CAB70005_

ALIGNMENTS

RESULT 1  
AY105508  
LOCUS AY105508 2180 bp mRNA linear HTC 16-OCT-2002  
DEFINITION Zea mays PCO114533 mRNA sequence.  
ACCESSION AY105508  
VERSION AY105508.1 GI:21208586  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 2180)  
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 2180)  
AUTHORS Coe,E.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES  
source Location/Qualifiers  
1..2180 /organism="Zea mays"



/mol\_type="mrna"  
/db\_xref="MaizeDB:636667"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Cornsensus Library"  
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores:  
Pred. No.:  
Score:  
Percent Similarity:  
Best Local Similarity:  
Query Match:  
DB:

1.06e-294  
3027.00  
99.66%  
99.33%  
74.28%  
3

Length:  
Matches:  
Conservative:  
Mismatches:  
Indels:  
Gaps:

2180  
591  
2  
2  
0  
0

US-10-051-909-32 (1-800) x AY105508 (1-2180)

QY	206	MetValPheGlyMetSerLeuSerProSerProAspTrpArgIleMetLeuGlyValLeu	225
Db	50	ATGGTGTGGGATGTCCTGTCGCCATCACCGGATTGGAGAAATTATGCTTGGTGTGCTC	109
QY	226	AlaIleProSerLeuPhePheGlyLeuThrIlePheTyrLeuProGluSerProArg	245
Db	110	GCGATACCTTCATTGTTCTCTTTGGTTTGACAAATATTTTATCTCTCTGAATCTCCAAGA	169
QY	246	TrpLeuValSerLysGlyArgMetAlaGluAlaLysLysValLeuGlnLysLeuArgGly	265
Db	170	TGGCTCGTTAGCAAGGTCGGATGGCAGAGGCAAAAAAGGTGTTGCAAAAGTTACGGTGG	229
QY	266	LysAspAspValSerGlyGluLeuSerLeuLeuGluGlyLeuGluValGlyGlyAsp	285
Db	230	AAAGACGATGTCACGGTGAATTGTCCCTTCCTCGAAGGTTGGAGGTTGGAGGAGAC	289
QY	286	ThrSerIleGluGluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThr	305
Db	290	ACTTCCATTGAAGAGTACATCATTTGGACCTGCCACCGAGGCAGCCGATGATCTTGTACT	349
QY	306	AspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAla	325
Db	350	GACGGTGATAAGGAACAAATCACACTTTATGGGCTGAAGAAAGGCCAGTCATGGATTGCT	409
QY	326	ArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGly	345
Db	410	CGACTTCTAAGGGACCCCATCATGCTTGGAAAGTGTCTTCTCTTGCATCTCGTCATGGG	469
QY	346	SerMetValAsnGlnSerValProLeuMetAspProIleValThrLeupheGlySerVal	365
Db	470	AGCATGGTGAACACAGAGTGATACCCCTTATGGATCCGATTGTGACACTTTTGGTAGTGTC	529
QY	366	HisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGly	385
Db	530	CATGAGAAATATGCCTCAAGCTGGAGGAAGTATGAGGAGCACATTGTTTCCAAACTTTGA	589
QY	386	SerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeu	405
Db	590	AGTATGTTCAGTGTACAGATCAGCATGCCAAAAATGAGCAGTGGGATGAAGAGAAATCTT	649
QY	406	HisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeu	425
Db	650	CATAGGGATGACGAGGAGTACGCACTGATGGTGCAGGAGGTGACTATGAGGACAAATCTC	709
QY	426	HisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHis	445
Db	710	CATAGCCCATTGCTGTCCAGGCAGGCAACAGGTGCGGAAGGGAAGGACATTGTGCACCAT	769
QY	446	GlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyAsp	465

Db	770	GGTCACCCGTGGAAGTGTCTTTAGCATGAGAAGGCAAGCCCTTTAGGGAGGGTGGAGAT	829
QY	466	GlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAlaTrpLysTrpSerGluLys	485
Db	830	GGTGTGAGCAGCACTGATATCGTGGGGGATGGCAGCTTGCTTGGAAATGGTCAGAGAA	889
QY	486	GluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGly	505
Db	890	GAAGGTGAGAAATGGTAGAAAGGAAGTGGTTTCAAAAGAGTCTACTTGCACCAAGAGGA	949
QY	506	ValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGlu	525
Db	950	GTTCTCTGGCTCAAGAAGGGGCTCAATTGTTTCACTTCCCGGTGGTGGCGATGTTCTTGAG	1009
QY	526	GlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGly	545
Db	1010	GGTAGTGAGTTGTACATGCTGCTGCTTTAGTAAGTCAGTCAGCACTTTTCTCAAAGGGT	1069
QY	546	LeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLys	565
Db	1070	CTTGCTGAACCACGCATGTCAGATGCTGCCATGGTTCACCCATCTGAGGTAGTGCCAAA	1129
QY	566	GlySerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyVal	585
Db	1130	GGTTCACGTTGAAAAAGATTGTTTGAACCTGGAGTGGCGTGCCCTGTTAGTCGGTGT	1189
QY	586	GlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGln	605
Db	1190	GGAATTTCAGATCCTTCAACAGTTTGTGGAAATAAACCGGTGTTCTGTACTATACCCACAA	1249
QY	606	IleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSer	625
Db	1250	ATTCTTGAAGCAAGCTGGTGTGGCAGTTATCTTTCCAAATTTGGTCTCAGCTCGGCATCA	1309
QY	626	AlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAla	645
Db	1310	GCATCCATCTTGATCAGTTCTCTCATACTTACTTAATGCTTCTTGCATTGGCTTTGCC	1369
QY	646	MetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeu	665
Db	1370	ATGCTGCTTATGGATCTTTCCGGAAGAAGGTTTTTGTGCTAGGCACAAATTCCAATCTG	1429
QY	666	IleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHis	685
Db	1430	ATAGCATCTCTAGTTATCTGTTGTGTCCAAATCTAATTGATTGGGTACACTAGCCCAT	1489
QY	686	AlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGly	705
Db	1490	GCTTTGCTCTCCACCATCAGTGTATCGTCTACTTCTGCTGCTTCGTTATGGGATTTGCT	1549
QY	706	ProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIle	725
Db	1550	CCCATCCCCAACATTTATGTGCAGAGATCTTCCAAACCAGGGTTCGTGGCCTCTGTATT	1609
QY	726	AlaIleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProVal	745
Db	1610	GCCATTTGTGCCTTTACATTTCTGGATCGGAGATATCATCGTCACCTACAGCCTTCCTGTG	1669
QY	746	MetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIle	765
Db	1670	ATGCTGAATGCTATTGGACTGGCGGGTGTTTTCAGCATATATGCAGTCGTATGCTTGATT	1729
QY	766	SerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIle	785
Db	1730	TCCTTTGTGTTCTGCTTCTTAAAGGTCCCTGAGACAAAGGGATGCCCTTGAGGTTATT	1789
QY	786	ThrGluPhePheAlaValGlyAlaLysGlnAlaAlaAlaLysAla	800
Db	1790	ACCGAATTCTTTGCAGTTGGTGCAGAGCAAGCGGTGCAAAAGCC	1834
RESULT 2			
CL980871			
LOCUS			
CL980871			
2292 bp			
DNA			
linear			
GSS 21-SEP-2004			



Db	1456	 GCGGCGGCGCTGGTGAGCCGGTCGATGCTGTACACCAAGGACGTCCTGTATCGGGCAGAGC	1515
Qy	551	MetSerAspAlaMetValHisProSerGlu-----ValAlaAlaLysGly	566
Db	1516	CGACCGAGCCGGCGTTTCGGAAACCCGCCGAGGGGTGGCGGCGGCGTCGACGGGC	1575
Qy	567	SerArgTrpLysAspLeuPheGluProGlyValArgArgAlaLeuLeuValGlyValGly	586
Db	1576	CGGGCGTGGCGGAGCTGCTCGAGCCGGGGCTCCGCCACGCCCTCTTCTCGGGCGTCAAC	1635
Qy	587	IleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle	606
Db	1636	ATCCAGATCCTCCAGCAATTCTCCGGCATCAACGGCGTGCTCTACTACACCCCGCAGATC	1695
Qy	607	LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla	626
Db	1696	CTCGACCGCGCGTCAGCGTCTCTCCGCCAGGCTCGGCCTCTCCGGCGACTCCACC	1755
Qy	627	SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet	646
Db	1756	TCCATCTCATCAGCGGCTCACCACGCTCCTCTCATGTCTCCGTCATCGGGCGTCGCATG	1815
Qy	647	LeuLeuMetAspLeuSerGlyArgPheLeuLeuLeuGlyThrIleProIleLeuIle	666
Db	1816	CGCCTCATGGACGCCCTCGGGCCGCCGCCCTCTCTCTGACGCTGCCCGTCTCGCTGC	1875
Qy	667	AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla	686
Db	1876	CGGTCCCTCGCCGTGCTCGTGGTGGCGAACGTCGTGCCCATGGCGGCGACCGCGACGCG	1935
Qy	687	LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro	706
Db	1936	CGCGTGTGACCGGGAGCGTCATCGTCTACTTCTGTGCTTCGTATGGGGTTCGGGCCCC	1995
Qy	707	IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla	726
Db	1996	ATCCCAACATCTCTGCGCCGAGATCTTCCCGACGAGGTGAGGGGACTCTGCATTGCC	2055
Qy	727	IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet	746
Db	2056	ATCTGCTCGCTGACGTTCTGGCTCGGCGACATCGCCGTCACGTACAGCCTCCCGTCATG	2115
Qy	747	LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeuIleSer	766
Db	2116	CTCAGCTCCGTGGGGCTCGCCGGCGTGTCTCTTCTACGCCGCCCGTGTGTCGCTCGCG	2175
Qy	767	PheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGluValIleThr	786
Db	2176	CTCGTGTTCGTGGCGCTCAAGGTGCCCGAGACCAAGGCCCTCCCGCTCGAGGTCATCATC	2235
Qy	787	GluPhePheAlaValGlyAlaLys	794
Db	2236	GAGTCTTCAACGTCGCGGCCAAG	2259

RESULT 3	
CL968379	
LOCUS	1932 bp DNA linear GSS 21-SEP-2004
DEFINITION	OsIFCC040330 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.
ACCESSION	CL968379
VERSION	CL968379.1 GI:52391393
KEYWORDS	GSS.
SOURCE	Oryza sativa (indica cultivar-group)
ORGANISM	Oryza sativa (indica cultivar-group)
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
	1 (bases 1 to 1932)
REFERENCE	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.
AUTHORS	An analysis of transcriptional regulation of the rice genome and
TITLE	

its comparison to Arabidopsis  
Unpublished (2004)  
COMMENT  
Contact: Chen Chen  
Department of Bioinformatic  
Beijing Institute of Genomics  
Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Fax: 86-10-80488676  
Email: chenchen@genomics.org.cn  
Rice genomic sequence.  
Class: exon-trapped.  
Location/Qualifiers  
1. .1932  
/organism="Oryza sativa (indica cultivar-group)"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:39946"  
/clone\_lib="Oryza sativa Express Library"  
/note="Oryza sativa exon trapped genomic sequences"

its comparison to Arabidopsis  
Unpublished (2004)  
Contact: Chen Chen  
Department of Bioinformatic  
Beijing Institute of Genomics  
Chinese Academy of Sciences, Beijing 101300, China  
Tel: 86-10-80481559  
Fax: 86-10-80488676  
Email: [chenchen@genomics.org.cn](mailto:chenchen@genomics.org.cn)  
Rice genomic sequence.  
Class: exon-trapped.





ORIGIN									
CoT selected genomic DNA library"									
Alignment Scores:									
Pred. No.:		1.78e-130		Length:		911			
Score:		1399.00		Matches:		267			
Percent Similarity:		98.90%		Conservative:		2			
Best Local Similarity:		98.16%		Mismatches:		3			
Query Match:		34.33%		Indels:		0			
DB:		8		Gaps:		0			
US-10-051-909-32 (1-800) x CC423708 (1-911)									
Qy	270	SerGlyGluLeuSerLeuLeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGlu	289						
Db	818	ACAGGTGAATTGTCCTTCTCTCGAAGGGTTGGAGGTTGGAGGAGACACTTCCATTGAA	759						
Qy	290	GluTyrIleIleGlyProAlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLys	309						
Db	758	GAGTACATCATTTGGACCTGCCACCGAGGAGCCGATGATCTTGTTTACTGACGGTGATAAG	699						
Qy	310	GluGlnIleThrLeuTyrGlyProGluGluGlyGlnSerTrpIleAlaArgProSerLys	329						
Db	698	GAACAAATCACACTTATGGCCCTGAAGAAGCCAGTCATGGATTGCTCGACCTTCCAAG	639						
Qy	330	GlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHisGlySerMetValAsn	349						
Db	638	GGACCCAGCATGCTTGGAAAGTGTGCTTCTCTGTCATCTCGTCATGGGAGCATGGTGAAC	579						
Qy	350	GlnSerValProLeuMetAspProIleValThrLeuPheGlySerValHisGluAsnMet	369						
Db	578	CAGAGTGTAACCCCTTATGGATCCGATTGTGACACTTTTGGTAGTGTCCATGAGAATATG	519						
Qy	370	ProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPheGlySerMetPheSer	389						
Db	518	CCTCAAGCTGGAGAAAGTATGAGGAGCACATGTTTCCAACTTTTGGAAAGTATGTTCACT	459						
Qy	390	ValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsnLeuHisArgAspAsp	409						
Db	458	GTCACAGATCAGCATGCCAAAATGAGCAGTGGGATGAAGAGAAATCTTCATAGGGATGAC	399						
Qy	410	GluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsnLeuHisSerProLeu	429						
Db	398	GAGGAGTACGCATCTGATGGTGTCAAGAGGTGACTATGAGGACAACTCTCCATAGCCCATG	339						
Qy	430	LeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHisHisGlyHisArgGly	449						
Db	338	CTGTCCAGGCAGGCAACAGGTGCGGAAGGGAAGGACATTTGTGCACCATGGTCACCGTGA	279						
Qy	450	SerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGlyAspGlyValSerSer	469						
Db	278	AGTGCTTTGAGCATGAGAAGGCAAGCCCTTAGGGGAGGGTGGAGATGGTGTGAGCAGC	219						
Qy	470	ThrAspIleGlyGlyTyrGlnLeuAlaTyrLysTrpSerGluLysGluGlyGluAsn	489						
Db	218	ACTGATATCGGTGGGGATGGCAGCTTGTCTTGGAAATGGTCAGAGAAGGAAGGTGAGAAT	159						
Qy	490	GlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyValProGlySer	509						
Db	158	GGTAGAAAGGAAGGTGGTTTCAAAGAGTCTACTTGCAACCAAGAGGGAGTTCTCTGGCTCA	99						
Qy	510	ArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPheGluGlySerGluPhe	529						
Db	98	AGAAGGGGCTCAATTGTTTTCACCTTCCCGGTGGTGGCGATGTTCTTTGAGGGTAGTGAGTTT	39						
Qy	530	ValHisAlaAlaAlaLeuValSerGlnSerAlaLeu	541						
Db	38	GTACATGCTGCTGCTTTAGTAAGTCAGTCAGCACTT	3						
RESULT 5									
BZ723844									
LOCUS									
DEFINITION BZ723844 PUCFI60TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTal133J23, linear GSS 24-FEB-2003									

genomic survey sequence.

ACCESSION  
BZ723844  
VERSION  
BZ723844.1  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
1 (bases 1 to 845)  
AUTHORS  
Whitelaw,C.A.; Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.  
TITLE  
Maize Genomics Consortium  
JOURNAL  
Unpublished (2003)  
COMMENT  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TF  
Class: sheared ends.  
Location/Qualifiers  
1..845  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBTal33J23"  
/clone\_lib="ZM\_0.6\_1.0\_KB"  
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
CoT selected genomic DNA library"

ORIGIN  
Alignment Scores:  
Pred. No.: 7.4e-129 Length: 845  
Score: 1382.50 Matches: 274  
Percent Similarity: 97.52% Conservative: 1  
Best Local Similarity: 97.16% Mismatches: 6  
Query Match: 33.93% Indels: 3  
DB: 8 Gaps: 0  
US-10-051-909-32 (1-800) x BZ723844 (1-845)  
QY 303 LeuValThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlnSer 322  
Db 3 CTTGTACTGACGGTGATAAGGAACAAATCACACTTTATGGCCTGAAGAAGGCCAGTCA 62  
QY 323 TrpIleAlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSer 342  
Db 63 TGGATTGCTCGACCTTCCAAGGGACCCAGCATGCTTGGAAAGTGTGCTTTCTCTTGCACT 122  
QY 343 ArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPhe 362  
Db 123 CGTCATGGGAGCATGGTGAAACCAGAGTGTAACCCCTTATGGATCCGATTGTGACACTTTT 182  
QY 363 GlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPhePro 382  
Db 183 GGTAGTGTCCATGAGAAATATGCCTCAAGCTGGAGGAAGTATGAGGAGCACATTGTTTCCA 242  
QY 383 AsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGlu 402  
Db 243 AACTTTGGAAGTATGTTCAAGTGTCAAGATCAGCATGCTAAAAATGAGCAGTGGGATGAA 302  
QY 403 GluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGlu 422  
Db 303 GAGAATCTTCATAGGGATGACGAGGAGTACGCATCTGTATGGTGCAGGAGGTGACTATGAG 362  
QY 423 AspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIle 442  
Db 363 GACAATCTCCATAGCCCATTTGCTGTCCAGGCAGGCAACAGGTGCGGAAGGAGGACATT 422





REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 1135)  
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
Unpublished  
2 (bases 1 to 1135)  
Genoscope.  
Direct Submission  
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full length  
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source  
1. .1135  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSLT5IL50ZE01"  
/tissue\_type="Silique"  
/plasmid="pCMVSPORT 6"  
complement(1. .1135)  
/gene="At4g35300"

ORIGIN

Alignment Scores:  
Pred. No.: 4,14e-117 Length: 1135  
Score: 1268.50 Matches: 247  
Percent Similarity: 82.82% Conservative: 47  
Best Local Similarity: 69.58% Mismatches: 58  
Query Match: 31.13% Indels: 3  
DB: 3 Gaps: 3  
US-10-051-909-32 (1-800) x CNS0A3H4 (1-1135)

QY 445 HisGlyHisArgGlySerAlaLeuSerMetArgArg---GlnThrLeuLeuGlyGluGly 463  
Db 1 CACACTGCTCATGGAACCTCTTTCTACCTTCAGACATGGAAGTCAAGTGCAGGAGCTCAA 60  
QY 464 GlyAspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTyrLysTrpSer 483  
Db 61 GGGGAAGGAGCGGGTAGTATGGGATTGGAGGTGGATGGCAAGTGGCATGGAATGGACG 120  
QY 484 GluLysGluGlyGluAsnGlyArgLysGluGlyGlyPhelysArgValTyrLeuHisGln 503  
Db 121 GAAAGAGAAGATGAATCGGACAGAAAGAGTGGGTTTAAACGGATATACTTCATCAA 180  
QY 504 GluGlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspVal 523  
Db 181 GAAGGTTTCCAGGATCTCGACGCTGGCTCAATTGTTTCATTGCCTGGTGGTATGGAACC 240  
QY 524 PheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSer 543  
Db 241 GGTGAG---GCAGATTTTGTACAGCGTCTGCTTTGGTTAGCCAAACCAGCTCTTTATTCC 297  
QY 544 LysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAla 563

Db 298 AAAGACCTTCTCAAAGAACATACAAATTGGTCTCTATGGTACATCCATCCGAA---ACA 354  
QY 564 AlaLysGlySerArgTyrLysAspLeuPheGluProGlyValArgArgAlaLeuVal 583  
Db 355 ACTAAAGGGTCAATTGGCATGATCTTCATGATCCTGGAGTCAAGCGTGCATTAGTCGT 414  
QY 584 GlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyValLeuTyrThr 603  
Db 415 GGAGTTGGACTTCAAATACTTCAGCAGTTCTCAGGCATCAACGGAGTCTTTACTACACA 474  
QY 604 ProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSer 623  
Db 475 CCGCAATCCTTGAGCAGCGGGTGTCTGGGATCCTACTATCGAACATGGGATTAGTTCT 534  
QY 624 AlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGly 643  
Db 535 TCCTCAGCATCCTTACTTATAAGTGCATTGACAAACCTTTGTGATGTTACCTGCAATAGCT 594  
QY 644 PheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIlePro 663  
Db 595 GTTGCAATGAGGCTCATGGATCTTTCTGGTCGAAAGGACCTTGCTTCCACCAGATACCA 654  
QY 664 IleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeu 683  
Db 655 ATCTCATAGCATCTCTATTGGTTTGTAGTAATCTCAAAATCTTGTTACATGAACAGCAT 714  
QY 684 AlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGly 703  
Db 715 GTGCACGCGGCTCTTATCAACCGTAAGCGTTTGCTCTACTTCTGCTTCTCGTGATGGGT 774  
QY 704 PheGlyProIleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeu 723  
Db 775 TTCTGCTCTGCTCCAAACATCCTCTGTTTCAGAGATTTTTCCAACTCGAGTCCGCGGAATC 834  
QY 724 CysIleAlaIleCysAlaPheThrPheTyrIleGlyAspIleValThrTyrSerLeu 743  
Db 835 TGCATCGCCATCTGCGCACTCACCTTCTGGATCTGTGACATAATCGTCACTTACAGTCTC 894  
QY 744 ProValMetLeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCys 763  
Db 895 CCGGTGCTGCTCAAATCCATCCATTGGACTAGCTGGTGTGTTTGGAAATGACGCAATCGTTGT 954  
QY 764 LeuIleSerPheValPheValPheLeuLysValProGluThrLysGlyMetProLeuGlu 783  
Db 955 TGCATTTTCATGGGCTCTTGTGTTTCATTAAAGTCCCGAAACTAAAGGCATGCCACTTGAA 1014  
QY 784 ValIleThrGluPheAlaValGlyAlaLysGlnAlaAlaAla 798  
Db 1015 GTCATCACAGAGTTCTTTCTGTTGGAGCTAGACAAAGCTGAAGCT 1059

RESULT 8  
CA065405  
LOCUS  
DEFINITION  
5', mRNA sequence.  
CA065405  
CA065405.1 GI:34916929  
EST.  
SOURCE  
ORGANISM  
Saccharum officinarum  
Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.  
1 (bases 1 to 833)  
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br  
Plate: 039 row: G column: 07  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers  
1. .833  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCACAD1039G07"  
/lab\_host="DH10B"  
/clone\_lib="AD1"  
/note="Organ: seedlings inoculated with Gluconacetobacter diazotrophicans; Vector: pSport1; Site\_1: Sali; Site\_2: Noti; An unidirectional cDNA library generated from [seedlings inoculated with Gluconacetobacter diazotrophicans]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"

FEATURES  
source

ORIGIN

Alignment Scores:  
Pred. No.: 3.77e-107 Length: 833  
Score: 1168.00 Matches: 244  
Percent Similarity: 90.65% Conservative: 8  
Best Local Similarity: 87.77% Mismatches: 26  
Query Match: 28.66% Indels: 2  
DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA065405 (1-833)

Qy 487 GlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGluGlyVal 506  
Db 2 GGTGAGGATGGTAAGAAGGAGGTGGTTTCAAAAAGAGTCTACTTACACCAAGAGGGAGTT 61  
Qy 507 ProGlySerArgGlySerIleValSerLeuProGlyGlyGlyAepValPheGluGly 526  
Db 62 CCTGGCTCAAGATGGGCTCAATTGTTTCACCTCTCGTGGTGGCGATGTTCTCGAGG 121  
Qy 527 SerGluPheValHisAlaAlaLeuValSerGlnSerAlaLeuPheSerLysGlyLeu 546  
Db 122 GCGAGTTTGATACATGCTGCTGCTTTAGTAAGCCAGTCAGCACCTTCTCGAAGGATCTT 181  
Qy 547 AlaGluProArgMetSerAspAlaAlaMetValHisProSerGluValAlaAlaLysGly 566  
Db 182 ACCGAACCAACGATGCTCGTGCTGCCATGGTTTACCCTCCGAGGTAGCTGCCAAGGT 241  
Qy 567 SerArgTrpLysAspLeuPheGluProGlyValArgAlaLeuLeuValGlyValGly 586  
Db 242 TCAAGTTGGAAAGATTGTTTGAACCTGGTGTGAGGCGTGCCCTGTAGTCGGTGTGGA 301  
Qy 587 IleGlnIleLeuGlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIle 606  
Db 302 ATTCAGATCCTTCAACAGTTTGCTGGAATAAATGGTGTCTGTACTATACCCCAAAATT 361  
Qy 607 LeuGluGlnAlaGlyValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAla 626  
Db 362 CTCGAGCAAGCTGGCGTGGCAGTTCTTCTTCCAATCTTGGTCTCAGCTCGGCATCAGCA 421  
Qy 627 SerIleLeuIleSerSerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMet 646  
Db 422 TCCATCTTGATCAGTTCTCTCACTACCTTACTGATGCTTCCCTAGCATGGCTTAGCCATG 481  
Qy 647 LeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIle 666

Db 482 AGACTTATGGATCTTTCTGGAAGAAGGTTTTTGCTGCTAGGCACAATTCCAATCTTGATA 541  
Qy 667 AlaSerLeuValIleLeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAla 686  
Db 542 GCATCTTTAGTTATCTGGTGGTGTCCAATGTTATTGACTTGGGTACAGTGGCCCATGCT 601  
Qy 687 LeuLeuSerThrValSerValIleValTyrPheCysCysPheValMetGlyPheGlyPro 706  
Db 602 GCGCTCTCCACAGTCAGTGTCAATCATCTACTTCTGCTGCTNTGTATGGGATTGGTCCC 661  
Qy 707 IleProAsnIleLeuCysAlaGluIlePheProThrArgValArgGlyLeuCysIleAla 726  
Db 662 ATCCCCAACATTTCTATGTGCAGAGATCTTNTCCACTAGGGTTCGCGGTCTCTGCATTGCC 721  
Qy 727 IleCysAlaPheThrPheTrpIleGlyAspIleIleValThrTyrSerLeuProValMet 746  
Db 722 ATCTGTGCCTTGACATTTNTNNGATCGAGACATCATTTGTACACCTACAGCCCTCCTGTGATG 781  
Qy 747 LeuAsnAlaIleGlyLeuAlaGlyValPheSerIleTyrAlaValValCysLeu 764  
Db 782 CTGAATGCTATTGGGACTAGCAGTGTTTTTCG-ATATATGCAGC-GTATGCTTG 833

RESULT 9  
CA146057

LOCUS SCVPRT2076D10.g RT2 Saccharum officinarum cDNA clone SCVPRT2076D10  
DEFINITION 5', mRNA sequence.

ACCESSION CA146057.1 GI:35044714  
VERSION CA146057  
KEYWORDS EST.  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.

REFERENCE 1 (bases 1 to 733)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br

Plate: 076 row: D column: 10  
Seq primer: T7 Promoter Primer.

FEATURES  
source

Location/Qualifiers  
1. .733  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCVPRT2076D10"  
/lab\_host="DH10B"  
/clone\_lib="RT2"  
/note="Organ: Root tips (0.3cm-long) from adult plants; Vector: pSport1; Site\_1: Sali; Site\_2: Noti; An unidirectional cDNA library generated from [Root tips (0.3cm-long) from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:  
Pred. No.: 8.04e-106 Length: 733  
Score: 1154.00 Matches: 221  
Percent Similarity: 93.88% Conservative: 9  
Best Local Similarity: 90.20% Mismatches: 14  
Query Match: 28.32% Indels: 1  
DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA146057 (1-733)

QY 340 LeuAlaSerArgHisGlySerMetValAsnGlnSerValProLeuMetAspProIleVal 359  
Db 1 CTCGCATCTCGTCATGGCAGCATGGTGAACCAAGAGTGTACCCCTTATGGATCCGATGTG 60  
QY 360 ThrLeuPheGlySerValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThr 379  
Db 61 ACACCTTTTGGGAGTGTCCACGAGAATATGGCTCAAGCTGGAGGAAGTATGAGGAGCACA 120  
QY 380 LeuPheProAsnPheGlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGln 399  
Db 121 TTGTTTCCAAACTTTGGAAGTATGTTTCAGTGTCCACAGATCAGCATGCCAAAATGAGCAG 180  
QY 400 TrpAspGluGluAsnLeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGly 419  
Db 181 TGGGACGAAGAGAAATCTTCACAGGGACGATGAGGAGTATGCATCTGATGGTGCAGAGGT 240  
QY 420 AspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGly 439  
Db 241 GATTATGAGGACAACTCCACAGCCCATGCTGTCCAGGCAGACAACAAGTGTGGAAGGA 300  
QY 440 LysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeu 459  
Db 301 AAGGACATTGTGCACCATGGTTCACCGTGGAAAGTCTTTGAGCATGAGAAGGCAAGCCTC 360  
QY 460 LeuGlyGluGlyGlyAspGlyValSerSerThrAspIleGlyGlyGlyTrpGlnLeuAla 479  
Db 361 TTGGGGGAGGCTGGAGAGGTGTGAGCAGCACTGATATTGGTGGAGGATGGCAGCTGCA 420  
QY 480 TrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPheLysArgVal 499  
Db 421 TGGAAATGGTCAGAGAAAGGAAGGTGAGGATGGTAAGAGGAAGGTGGTCTCAAAAGAGTC 480  
QY 500 TyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGly 519  
Db 481 TACTTACACCAAGAGGGAGTTCTCTGGCTCAAGAAATGGGCTCAATTGTTTCACTTCTGGT 540  
QY 520 GlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSer 539  
Db 541 GGTGGCAATGTTCTCTGAGGTGGCGAGTGTGTACATGCTGCTGCTTTAGTAAGCCGTC 600  
QY 540 AlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisPro 559  
Db 601 GCACCTTTCTTCGAAGATCTTACCGAACCAACGCGCATGCTGGTGTGCCATGGTTCAACCCA 660  
QY 560 SerGluValAlaAlaLysGlySerArg-TrpLysAspLeuPheGluProGlyValArgAr 579  
Db 661 TCCAGGTAGCTGCCAAAGGTTCAAAGTTGGAAGATTGTTTGAACCTGGTGTGAGGCG 720  
QY 579 gAlaLeuLeuVal 583  
Db 721 TGCCCTTGTAGTC 733

RESULT 10  
CA075335  
LOCUS SCJFAM1069A06.g AM1 Saccharum officinarum cdna clone SCJFAM1069A06  
DEFINITION 696 bp mRNA linear EST 23-SEP-2003  
5', mRNA sequence.  
ACCESSION CA075335  
VERSION CA075335.1 GI:34927608  
KEYWORDS EST.  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
complex.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 696)  
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089

Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: 069 row: A column: 06  
Seq primer: T7 Promoter Primer.

FEATURES  
source

Location/Qualifiers  
1..696

/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCJFAM1069A06"  
/lab\_host="DH10B"  
/clone\_lib="AM1"

/note="Organ: Apical meristem and tissues surrounding of  
mature plants; Vector: pSport1; Site\_1: SalI; Site\_2:  
NotI; An unidirectional cDNA library generated from  
[Apical meristem and tissues surrounding of mature  
plants]. cDNA was prepared from polyA+ mRNA using  
SuperScript plasmid System Kit (Invitrogen). The  
double-strand cDNAs were fractionated in a sepharose  
CL-2B 40cm-columns and fragments sizing between 0.8 and  
1.5 Kb were directionally cloned into the vector. Details  
of each source of RNA and library construction can be  
obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:  
Pred. No.: 1.21e-104 Length: 696  
Score: 1142.00 Matches: 214  
Percent Similarity: 95.22% Conservative: 5  
Best Local Similarity: 93.04% Mismatches: 11  
Query Match: 28.02% Indels: 0  
DB: 6 Gaps: 0

US-10-051-909-32 (1-800) x CA075335 (1-696)

QY 305 ThrAspGlyAspLysGluGlnIleThrLeuTyrGlyProGluGluGlnSerTrpIle 324  
Db 1 ACTGATGGTGATAAGGAACAAATCACACTTTATGGGCCTGAAGAGGCCAGTCATGGATT 60  
QY 325 AlaArgProSerLysGlyProIleMetLeuGlySerValLeuSerLeuAlaSerArgHis 344  
Db 61 GCTCGACCTTCCAAGGGACCCAGCATGCTTGGAAAGTGTCTTGCTCTCGCATCTCGTCAT 120  
QY 345 GlySerMetValAsnGlnSerValProLeuMetAspProIleValThrLeuPheGlySer 364  
Db 121 GGCAGCATGGTGAACCAAGAGTGTACCCCTTATGGATCCGATTGTGACACTTTTGGGAGT 180  
QY 365 ValHisGluAsnMetProGlnAlaGlyGlySerMetArgSerThrLeuPheProAsnPhe 384  
Db 181 GTCCACGAGAAATATGCCTCAAGCTGGAGGAAGTATGAGGAGCACATTGTTCCAAACTTT 240  
QY 385 GlySerMetPheSerValThrAspGlnHisAlaLysAsnGluGlnTrpAspGluGluAsn 404  
Db 241 GGAAGTATGTTTCAGTGTACAGATCAGCATGCCAAAAAATGAGCAGTGGGACGAAGAAT 300  
QY 405 LeuHisArgAspAspGluGluTyrAlaSerAspGlyAlaGlyGlyAspTyrGluAspAsn 424  
Db 301 CTTCACAGGGACGATGAGGAGTATGCATCTCATGTTGTCAGGAGGTGATATGAGGACAAT 360



Qy

425

LeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGlyLysAspIleValHis

444

Db

361

CTCCACAGCCCATGTGTCAGGCAGACAACAAGTATGGAAGGAAGGACATTGTGCAC

420

Qy

445

HisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeuLeuGlyGluGlyGly

464

Db

421

CATGGTCACCGTGGAAGTGCTTTGAGCATGAGAAGGCAAAGCCTCTTGGGGAGGCTGGA

480

Qy

465

AspGlyValSerSerThrAspIleGlyGlyGlyTyrGlnLeuAlaTrpLysTrpSerGlu

484

Db

481

GAGGGTGTGAGCAGCACTGATATTGGTGGAGGATGGCAGCTTGCAATGGAAATGGTCAGAG

540

Qy

485

LysGlyGlyGluAsnGlyArgLysGluGlyGlyPheLysArgValTyrLeuHisGlnGlu

504

Db

541

AAGGAAGGTGAGGATGGTAAGAAGGAAGGTGGTTTCAAAAGAGTCTACTTACACCAAGAG

600

Qy

505

GlyValProGlySerArgArgGlySerIleValSerLeuProGlyGlyGlyAspValPhe

524

Db

601

GGAGTTCCTGNCTCAAGAATGGGCTCANTNGTNTCACTTCCTGGTGGCGATGTTCC

660

Qy

525

GluGlySerGluPheValHisAlaAlaAla

534

Db

661

GAGGGTGGCGAGTTNGTACATGCTGCTGCT

690

RESULT 11

CK208781

LOCUS

DEFINITION

FGAS020500 Triticum aestivum FGAS: Library 5 GATE 7 Triticum aestivum cDNA, mRNA sequence.

ACCESSION

CK208781

VERSION

CK208781.1

GI:395711171

KEYWORDS

EST.

SOURCE

Triticum aestivum (bread wheat)

ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

1 (bases 1 to 1082)

Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D., Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A., Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D., Penniket,C., Roach,J.L. and Sarhan,F.

Functional Genomics of Abiotic Stress In Wheat and Canola Crops Unpublished (2003)

Contact: Wm L Crosby

Bioinformatics

University of Saskatchewan, Department of Computer Science 1C101 Engineering Building, 57 Campus Drive, Saskatoon, Saskatchewan, S7N 5A9, Canada

Tel: 306 966 1769

Fax: 306 966 2033

Email: fgas\_ests@cs.usask.ca

Phred (default parameters). It is the raw base calls. To aid in the identification of the high quality insert the software Lucy (default parameters) has been run on this sequence. Lucy identified the region [31,846].

Plate: L5B015

row: E

column: 19.

Location/Qualifiers

1. .1082

/organism="Triticum aestivum"

/mol type="mRNA"

/db\_xref="taxon:4565"

/clone lib="Triticum aestivum FGAS: Library 5 GATE 7"

/note="Vector: pCMV.SPORT6; Crown and developmental stages of spike formation in wheat cultivar Norstar. 4 mRNA populations were combined before constructing the library. The first mRNA population is from 1cm crown sections after 30 days of cold acclimation. The second is from 1cm crown sections after 11 days of deacclimation (before deacclimation plants were fully vernalized for 49 days).

The third is from different developmental stages of spike

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.:

4.15e-102

Length:

1082

Score:

1120.00

Matches:

242

Percent Similarity:

81.23%

Conservative:

35

Best Local Similarity:

70.97%

Mismatches:

49

Query Match:

27.48%

Indels:

20

DB:

7

Gaps:

4

US-10-051-909-32 (1-800) x CK208781 (1-1082)

Qy

56

MetSerGlyAlaValLeuValAlaIleValAlaSerIleGlyAsnLeuLeuGlnGlyTyr

75

Db

65

ATGTCGGCGCCGTGCTGGTCCCATAGCGGCTCCATCGGGAACCTGCTGCAGGGGTGG

124

Qy

76

AspAsnAlaThrIleAlaAlaValLeuTyrIleLysGluPheGlnLeuGlnAsn

95

Db

125

GATAATGCGACCATTCAGGTGCGCTCCTGTACATAAAGAAGAGTTCAGCTTGGAGACC

184

Qy

96

GluProThrValGluGlyLeuIleValSerMetSerLeuIleGlyAlaThrIleValThr

115

Db

185

CAGCCCTTGATCGAGGCGCTCATCGTGCCCATGTGCTCATCGGGCGACGGTTATCACC

244

Qy

116

ThrPheSerGlyProLeuSerAspSerIleGlyArgArgProMetLeuIleLeuSerSer

135

Db

245

ACGTTCTCCGGGGCGGTGGCTGATGCTGTTGGTAGGGGCCCTGCTGATCGCCTCGTCT

304

Qy

136

IleLeuTyrPhePheSerGlyLeuIleMetLeuTyrSerProAsnValTyrValLeuLeu

155

Db

305

GTCTCTACTTTGTAGTGGCTGGTGATGCTCTTGGCGCCCAACGTCTATGTGTTGCTC

364

Qy

156

LeuAlaArgPheValAspGlyPheGlyIleGlyLeuAlaValThrLeuValProLeuTyr

175

Db

365

TTGGCCAGGCTCATCGACGGGTTCGGTATCGGTTTGGCTGTCAACCTTGTCCCTCTTTAC

424

Qy

176

IleSerGluIleAlaProSerGluIleArgGlyLeuLeuAsnThrLeuProGlnPheSer

195

Db

425

ATTTCCGAGACCGCCCCGACTGACATTAGAGGGCTGCTGAACACGCTGCCGAGTTTCAGT

484

Qy

196

GlySerGlyGlyMetPheLeuSerTyrCysMetValPheGlyMetSerLeuSerProSer

215

Db

485

GGGTcAGGAGGGATGTTCCCTTTCTTACTGCATGGTGTTCACCATGTCGCTCATGCCGAG

544

Qy

216

ProAspTrpArgIleMetLeuGlyValLeuAlaIleProSerLeuPhePheGlyLeu

235

Db

545

CCTGACTGGAGAAATCATGCTTGGGGTTTTTGTGATCCCGTTCGCTTATGTACTTTGCATTG

604

Qy

236

ThrIlePheTyrLeuProGluSerProArgTrpLeuValSerLysGlyArgMetAlaGlu

255

Db

605

ACTGTCTTCTATTGCCCCGAGTCGCCGAGATGGCTTGTGAGCAAAGGAAGATGGCCGAG

664

Qy

256

AlaLysLysValLeuGlnLysLeuArgGlyLysAspAspValSerGlyGluLeuSerLeu

275

Db

665

GCCNACGAGTGTTCGAAGGACTCGGGGAAGGAAGATGCTCAGGAGAAATGGCCCTT

724

Qy

276

LeuLeuGluGlyLeuGluValGlyGlyAspThrSerIleGluGluTyrIleIleGlyPro

295

Db

725

CTTGTGAAGGATTGGGTGTTGGGAAAGACACACATTTTGAAGAATACATAATTGGGCT

784

Qy

296

AlaThrGluAlaAlaAspAspLeuValThrAspGlyAspLysGluGlnIleThrLeuTyr

315

Db

785

GATGATGAGCTTGCTGATGACGGTCTGGCTCCA---GATGAAGAGAA-GTGAAGCTGTAC

840

Qy

316

GlyProGluGluGlyGlnSerTrpIleAlaArgProSerLysGlyProIleMetLeuGly

335

formation (5 to 50mm) that still have not emerged from the leaf (dissection required). The last is from different developmental stages of spike and seed formation after having emerged from the leaf (visible). First strand synthesis in this library was done in the presence of methylated dcrp thereby protecting from internal cleavage with NotI."



ACCESSION CA170109  
VERSION CA170109.1 GI:35093205  
KEYWORDS EST.  
SOURCE Saccharum officinarum  
ORGANISM Saccharum officinarum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.  
REFERENCE 1 (bases 1 to 871)  
AUTHORS Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
TITLE The libraries that made SUCEST  
JOURNAL Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
COMMENT Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br  
Plate: 083 row: B column: 09  
Seq primer: T7 Promoter Primer.  
FEATURES  
source  
1..871  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCQGSB1083B09"  
/lab\_host="DH10B"  
/clone\_lib="SB1"  
/note="Organ: Stalk Bark from adult plants; Vector: pSport1; Site\_1: SalI; Site\_2: NotI; An unidirectional cDNA library generated from [Stalk Bark from adult plants]. cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucest.lad.ic.unicamp.br/public"  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.11e-100 Length: 871  
Score: 1100.00 Matches: 215  
Percent Similarity: 86.11% Conservative: 33  
Best Local Similarity: 74.65% Mismatches: 40  
Query Match: 26.99% Indels: 2  
DB: 6 Gaps: 0  
US-10-051-909-32 (1-800) x CA170109 (1-871)  
Qy 420 AspTyrGluAspAsnLeuHisSerProLeuLeuSerArgGlnAlaThrGlyAlaGluGly 439  
Db 9 GACTATGAAACAATGTACACAGCCCTCTGCTGTCCGACAGATACAAAGTGCCGAAGG 68  
Qy 440 LysAspIleValHisHisGlyHisArgGlySerAlaLeuSerMetArgArgGlnThrLeu 459  
Db 69 AAGGGCATGTCACATCACGGGCACCGTGGAAAGTCTTTGGGTTTGAGAAGAAGCCTG 128  
Qy 460 LeuGlyGluGlyAspGlyValSerSerThrAspIleGlyGlyTrpGlnLeuAla 479  
Db 129 TCCGATGAGGGTGGTGAGGCAGCAACCAGCACTGGCATTTGGTGGCGATGCCGCTGCC 188  
Qy 480 TrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyPheLysArgVal 499  
Db 189 TGGAAATGGTCAGAGCGAGAAGGTGAGGACGGTAAGAGGAAGGAGTTTCAAAGAATC 248  
Qy 500 TyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSerLeuProGly 519

Db 249 TACTTACACCAAGAGGGAGTTTGCTGGCTCAAGAAGGGGATCTGTGTCTCACTTCCTCGT 308  
Qy 520 GlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuValSerGlnSer 539  
Db 309 GGAGGTGATGCCTCTGAAGGTGGTAAGTTCATACATGCTGCTGTCTGTGTCAGCCAGTCA 368  
Qy 540 AlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMetValHisPro 559  
Db 369 GCTCTTTACCCGAGGGATATTACAGAACAGCGCATGGCCGGTCCAGCTACAATGCACCCA 428  
Qy 560 SerGluValAlaAlaLysGlySerArgTrpLysAspLeuPheGluProGlyValArgArg 579  
Db 429 TCAGAGCAGCTGCCAAAGTTCCAAGCTGGAGAGACCTGTTTGAAACCTGGTGTGAGACGT 488  
Qy 580 AlaLeuValGlyValGlyIleGlnIleLeuGlnGlnPheAlaGlyIleAsnGlyVal 599  
Db 489 GCCCTGTGGTTGGTATTGGAATTCAGATCCTTCAACAGTTCGCGGGAATAAATGGNGTT 548  
Qy 600 LeuTyrTyrThrProGlnIleLeuGluGlnAlaGlyValAlaValIleLeuSerLysPhe 619  
Db 549 CTCTACTACACTCCGCAAATTTATGGAGCAAGCTGGTGTGGCAGTTCTTATTTCCAATCTT 608  
Qy 620 GlyLeuSerSerAlaSerAlaSerIleLeuIleSerSerLeuThrThrLeuLeuMetLeu 639  
Db 609 GGTCTCAGTTCGGCATCAGCATCCATCCTAATTAGTTCTGTCACCTGCCCTACTTATGCTN 668  
Qy 640 ProCysIleGlyPheAlaMetLeuLeuMetAspLeuSerGlyArgArgPheLeuLeuLeu 659  
Db 669 CCTAGCATTTGGTTTAGCCATGAGACTTATGGACGTATCTGGAAGAAGGTTTCTGTGCTA 728  
Qy 660 GlyThrIleProIleLeuIleAlaSerLeuValIleLeuValValSerAsnLeuIleAop 679  
Db 729 AGCACAATTCAGTCTTGATAGCATCTTTAATTGCTCTGGNTGTGTCTAATGTTATCGAG 788  
Qy 680 LeuGlyThrLeuAlaHisAlaLeuLeuSerThrValSerValIleValTyrPheCysCys 699  
Db 789 TTGGGTACAGTGGTGCCATGCTGTGCTCTCCACATT-AGTGTCACTCACCTACTTATGCTGC 847  
Qy 700 PheValMetGlyPheGlyProfile 707  
Db 848 TTCAA-AATGGCTTTGGCCCCCAT 870  
RESULT 14  
BX841312  
LOCUS BX841312  
DEFINITION Arabidopsis thaliana Flowers and buds Col-0 Arabidopsis thaliana cDNA clone GSLTFB38ZG07 SPRIM, mRNA sequence.  
ACCESSION BX841312  
VERSION BX841312.1 GI:42535395  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
REFERENCE 1 (bases 1 to 1107)  
AUTHORS Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C., Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V., Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.  
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation  
JOURNAL Unpublished (2004)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.





Pred. No.:		7.34e-97	Length:	848
Score:		1066.50	Matches:	215
Percent Similarity:		82.53%	Conservative:	26
Best Local Similarity:		73.63%	Mismatches:	36
Query Match:		26.17%	Indels:	15
DB:		7	Gaps:	4
US-10-051-909-32 (1-800) x CF445664 (1-848)				
Qy	477	GlnLeuAlaTrpLysTrpSerGluLysGluGlyGluAsnGlyArgLysGluGlyGlyPhe	496	
Db	1	CAGCTTGCTTGGAAAAAGAGTGAAGGGGAGGACAGAATGGTAAAAAGAAAGTGGGTTT	60	
Qy	497	LysArgValTyrLeuHisGlnGluGlyValProGlySerArgArgGlySerIleValSer	516	
Db	61	AAAAGGATATATTACAT--GAAGGGTTGCTGGTTCTAGAAGAGGTCATTTGTTCC	117	
Qy	517	LeuProGlyGlyGlyAspValPheGluGlySerGluPheValHisAlaAlaLeuVal	536	
Db	118	ATACAG---GGAGGTGATGTTCCGGAAGATAGCGAGTACGTGCAGGTGGCAGCTTTGGTT	174	
Qy	537	SerGlnSerAlaLeuPheSerLysGlyLeuAlaGluProArgMetSerAspAlaAlaMet	556	
Db	175	AGTGAGTCTGCTCTTTATTTCGAAG-----GATAGAAGTGGG	210	
Qy	557	ValHisPro-----SerGluValAlaAlaLysGlySerArgTrpLysAsp	571	
Db	211	CATCATCTGTGGCTATGGTTGATTCTAATGAAGCGAGTAAAGGGATGAGATGGTCGGAT	270	
Qy	572	LeuPheGluProGlyValArgAlaLeuLeuValGlyValGlyVileGlnIleLeuGln	591	
Db	271	CTTCTTGAACCAAGGTGTTAGGAGTGCTTTGATTTGTTGGAGTTGGGATTCAGATTCTCCAG	330	
Qy	592	GlnPheAlaGlyIleAsnGlyValLeuTyrTyrThrProGlnIleLeuGluGlnAlaGly	611	
Db	331	CAGTTTTCAGGCATCAACGGTGTTCTGTACTACACACCACAGATTCTTGAACAAGCAGGT	390	
Qy	612	ValAlaValIleLeuSerLysPheGlyLeuSerSerAlaSerAlaSerIleLeuIleSer	631	
Db	391	GTTGGAGTTTCTCCTCTCAAATCTCGGACTCACTGATTTCTTCAGCCTCCATTCTTATAAGT	450	
Qy	632	SerLeuThrThrLeuLeuMetLeuProCysIleGlyPheAlaMetLeuLeuMetAspLeu	651	
Db	451	GCATTGGTTACTCTACTAAATGCTTCCATCTATCGGTAATTGCTATGAGATTTCATGGACATA	510	
Qy	652	SerGlyArgArgPheLeuLeuLeuGlyThrIleProIleLeuIleAlaSerLeuValIle	671	
Db	511	GCTGGAAGAGGTTTCTTCTTCTATCCACAATACCCATCCTAATCTTATCTTTAGTCATA	570	
Qy	672	LeuValValSerAsnLeuIleAspLeuGlyThrLeuAlaHisAlaLeuLeuSerThrVal	691	
Db	571	CTGGTAATATCCAACATAATCGAGTTCGGCACAGTGGTACATGCTATCCTTTCAACTGCA	630	
Qy	692	SerValIleValTyrPheCysCysPheValMetGlyPheGlyProIleProAsnIleLeu	711	
Db	631	AGTGTCAATTGTATACTTCTGTGTTTTGTACGGGTTTTTGGACCCATACCCCAACATTCTA	690	
Qy	712	CysAlaGluIlePheProThrArgValArgGlyLeuCysIleAlaIleCysAlaPheThr	731	
Db	691	TGTGCTGAATATTCCCAACCAGAGTCAGAGGCATCTGCATTGCTATCTCGCTCTCACG	750	
Qy	732	PheTrpIleGlyAspIleIleValThrTyrSerLeuProValMetLeuAsnAlaIleGly	751	
Db	751	TTCTGGATCGGTGACATAATTGTACATAATACCCTTCCTCTGTATGCTCAACTCCATTGGT	810	
Qy	752	LeuAlaGlyValPheSerIleTyrAlaValValCys	763	
Db	811	CTGGCTGGGGTTTTTGGTATTTACGCTACTGTCTGT	846	

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 13:34:09 ; Search time 5504.66 Seconds  
(without alignments)  
4867.828 Million cell updates/sec

Title: US-10-051-909-36  
Perfect score: 2779  
Sequence: 1 PSSSSSFRPAGKKKKKNQG.....TIFVSLSIQRQLQWLPECLS 553

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074037\_14048/app\_query.fasta\_1.1678  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909 @CGN 1 1 9091 @runat\_13042005\_074037\_14048 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2133	76.8	1870	8 AK100792	AK100792 Oryza sat
2	2017.5	72.6	2266	8 AK065497	AK065497 Oryza sat
3	1821.5	65.5	1948	8 AK068594	AK068594 Oryza sat
4	1757.5	63.2	2020	8 BVU43629	U43629 Beta vulgar

5	1746.5	62.8	1557	8	AY093274	AY093274 Arabidops
6	1746.5	62.8	1866	8	AY059848	AY059848 Arabidops
7	1738.5	62.6	1695	8	AF412060	AF412060 Arabidops
8	1734.5	62.4	1464	8	AY124845	AY124845 Arabidops
9	1726.5	62.1	1914	8	AF462803	AF462803 Arabidops
10	1710	61.5	1377	6	AX653119	AX653119 Sequence
11	1693	60.9	1825	8	BT014234	BT014234 Lycopersi
12	1577	56.7	1428	6	AX653120	AX653120 Sequence
13	1453.5	52.3	204649	8	AC120986	AC120986 Oryza sat
c 14	1137.5	40.9	120977	8	AC025808	AC025808 Genomic s
15	1117.5	40.2	68041	8	AC025814	AC025814 Arabidops
c 16	992.5	35.7	99587	8	AC007858	AC007858 Oryza sat
c 17	982	35.3	156643	8	AC120988	AC120988 Oryza sat
18	965.5	34.7	1699	8	LES278765	AJ278765 Lycopersi
19	954.5	34.3	2008	8	AK106658	AK106658 Oryza sat
20	951.5	34.2	1741	8	AY084684	AY084684 Arabidops
21	944.5	34.0	1778	8	AY064144	AY064144 Arabidops
22	943.5	34.0	1435	8	BT000053	BT000053 Arabidops
23	943.5	34.0	1686	8	AY120715	AY120715 Arabidops
24	939	33.8	1449	8	BT000608	BT000608 Arabidops
25	935.5	33.7	1747	8	AY087180	AY087180 Arabidops
26	932	33.5	1806	8	BT008661	BT008661 Arabidops
27	932	33.5	1847	8	ATH249967	AJ249967 Arabidops
28	920.5	33.1	1473	6	AX507636	AX507636 Sequence
29	920.5	33.1	1678	8	D89051	D89051 Arabidops
30	899	32.3	1359	6	AX506932	AX506932 Sequence
31	898.5	32.3	1420	8	AY091216	AY091216 Arabidops
32	898.5	32.3	1660	8	AY063856	AY063856 Arabidops
33	891.5	32.1	1755	8	AK064996	AK064996 Oryza sat
34	878.5	31.6	1729	8	AF367260	AF367260 Arabidops
35	877.5	31.6	1407	6	AX506596	AX506596 Sequence
36	877.5	31.6	1407	6	AX651294	AX651294 Sequence
37	877.5	31.6	1413	8	AY133547	AY133547 Arabidops
38	870.5	31.3	1606	8	AY026255	AY026255 Arabidops
39	866	31.2	1578	8	ATH249968	AJ249968 Arabidops
40	858.5	30.9	1757	8	AY048207	AY048207 Arabidops
41	858.5	30.9	1757	8	AY123991	AY123991 Arabidops
42	853.5	30.7	1787	8	AY026254	AY026254 Arabidops
43	845	30.4	1668	8	AY058056	AY058056 Arabidops
44	843.5	30.4	1377	8	BT002699	BT002699 Arabidops
45	837.5	30.1	1209	6	AX653566	AX653566 Sequence

ALIGNMENTS

RESULT 1	AK100792	AK100792	Oryza sativa (japonica cultivar-group)	1870 bp	mRNA	linear	PLN 24-JUL-2003
LOCUS	AK100792	Oryza sativa (japonica cultivar-group)	cdna clone:J023121D20, full				
DEFINITION	AK100792	Oryza sativa (japonica cultivar-group)	insert sequence.				
ACCESSION	AK100792	AK100792	GI:32986001				
VERSION	AK100792.1	AK100792.1	GI:32986001				
KEYWORDS	FLI_CDNA; CAP trapper.	FLI_CDNA; CAP trapper.					
SOURCE	Oryza sativa (japonica cultivar-group)	Oryza sativa (japonica cultivar-group)					
ORGANISM	Oryza sativa (japonica cultivar-group)	Oryza sativa (japonica cultivar-group)					
REFERENCE	1	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.					
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tannoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Nariakawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN; Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,						

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 1870)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1..1870

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="J023121D20"

FEATURES

source

ORIGIN

Alignment Scores:  
Pred. No.: 3e-150 Length: 1870  
Score: 2133.00 Matches: 434  
Best Local Similarity: 87.77% Conservative: 47  
Best Local Similarity: 79.20% Mismatches: 51  
Query Match: 76.75% Indels: 17  
DB: 8 Gaps: 7

US-10-051-909-36 (1-553) x AK100792 (1-1870)

QY 1 ProSerSerSerSerSerPheArgProAlaGlyLysLysLysLysLysLysAsnGlnGly 20  
Db 66 CCTTCCCAATTCATCT-----CGTCCTCTCTCTCGTAACCTG----- 101  
QY 21 LeuArgArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArg-----Val 38  
Db 102 ---TAAAGAGATCTCGTCTCGCCTCGGGAATCCAACCATTCGCGGAGTAGAGATCGG 158  
QY 39 MetGlyGlyGlySerAsnArg---GlyGlyAlaGlyAlaGlyGluGluSerGlySerAsp 57  
Db 159 GAGGGCGGCGCCATGAACAGGGCGGCGGCGTCTGGCGTATGAGAGCGGAAGCGAC 218  
QY 58 HisAsp-----GlyValLeuArgArgProLeuLeu---AsnThrGlySerTrpTyrArg 74  
Db 219 TACGAGAGCGGCGGCGCATGCGGAAGCCGCTGCTGATGCACACGGGGAGCTGGTACAGG 278  
QY 75 MetSerSerArgGlnSerSerPhe---AlaProGlyThrSerSerMetAlaValLeuArg 93  
Db 279 ATGGGTTCGCGGCAGGGAGGAGCTCACCGGCGGGGACCTCGTCCATGGCCATCCTACGC 338  
QY 94 GluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPhe 113  
Db 339 GAGTCCCAAGTCTCCGCTTCTCTGCACGCTCATCGTCGCGCTCGGCCCATCCAGTTC 398  
QY 114 GlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeu 133  
Db 399 GGATTACAGGGGGGCTTCTCTCTCCCGACGACGAGACGCGCATCATCCGAGACCTCGACCTC 458  
QY 134 SerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAla 153  
Db 459 ACCCTCTCCGAGTTCTCGGTGTTCGGATCGTGTCCACAGTCGGCGCCATGGTGGGCG 518  
QY 154 IleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAla 173  
Db 519 ATTGCCAGTGGTCAGATGGCCGAGTACATTGGCGCGCAAGGGGTCAATTGATGATGCTGCA 578  
QY 174 IleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyr 193  
Db 579 ATTCCAAACATCATTTGGTTGGCTTGCCATCTCCTTTGAAAAGACTCATCGTTCTTTAT 638  
QY 194 MetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyr 213  
Db 639 ATGGGACGATTGCTCGAGGGGTTTGGTGTGGTGTCTATCTTATACGGTGCCAGTTTAC 698  
QY 214 IleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSer 233  
Db 699 ATAGCAGAAATATCACCTCAAAACATGAGAGGTGCTCTTGGCTCAGTGAATCAGTTATCT 758  
QY 234 ValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeu 253  
Db 759 GTAACCGTAGGTATATTGTTGGCATATTGCTCGGCATGTTGTTCTTGGAGGCTTCTT 818  
QY 254 AlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGlu 273  
Db 819 GCTGTAATAGGAATCTTGCCCTTGCACTGTGTGTGATACCTTGGCCTATTCTTCATCCAGAA 878  
QY 274 SerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnVal 293  
Db 879 TCCCCAAGATGGTTGGCAAAGATGAACATGATGGATGATTTTGGAGACTTCTTTTACAAGTT 938  
QY 294 LeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAla 313  
Db 939 CTGAGGGGATTTGAGACTGACATCAGCGCGGAAGTGAATGATATATAAAGAGAGCAGTAGCG 998

QY 314 SerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThr 333  
||||:|||||  
Db 999 TCAGCAACAAGGACAACGATCCGTTTTCAAGAATTAAACCAGAGAANAATACCGCAC 1058  
  
QY 334 ProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyLeAsnGlyVal 353  
|||||:|||||  
Db 1059 CCCCTAATACTAGGAATTGGCCTACTTGTACTGCAACAGCTAAGTGGATCAATGGAA 1118  
  
QY 354 LeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThr 373  
|||||  
Db 1119 TTGTTTTATGCAGGTAGCATCTTCAAAGCAGCAGGTCTCACAAACAGTGACTTGGCTACA 1178  
  
QY 374 CysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArg 393  
||||:|||||  
Db 1179 TGTGCACCTTGGTGCTATCCAGGTTCTTGTACAGGAGTTACAACCTGGTTATTAGACAGA 1238  
  
QY 394 AlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAla 413  
|||||  
Db 1239 GCTGGCCGACGATCCTCTTATCATCTCTTCTGTCTGGGATGACTCTAAGCCTCCTTGCA 1298  
  
QY 414 ValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIle 433  
||||:|||||  
Db 1299 GTTGCTGTTGTATTTTCTCAAGGATAGCATTTCAAGAATTCTCACATGTACTACACC 1358  
  
QY 434 LeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet 453  
||||:|||||  
Db 1359 TTAAGTATGATCTCCTTGGTTGCTCTTGTGGCTTTTGTAAATCGCCTTCTCCTTCGGTATG 1418  
  
QY 454 GlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly 473  
|||||  
Db 1419 GGTGCCATTCCATGGATCATATGTTCAGAGATCCTCCCGGTTAGTATCAAGAGTCTCGCA 1478  
  
QY 474 GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsn 493  
|||||  
Db 1479 GGAAGCTTTGCGACGCTCGCCAACTGGCTTACATCCTTTTGGAAATAACAATGACAGCAAAC 1538  
  
QY 494 LeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPhe 513  
|||||  
Db 1539 TTGATGCTTAGCTGGAGTGTCTGGAGGACCTTTGTGTCTCTACATGTCGTGAGTGTCTTC 1598  
  
QY 514 ThrIleValPheValValLeuTrpValProGluThrLysGly\*\*AsnSerArgGlyAsp 533  
||||:|||||  
Db 1599 ACCCTCGTGTGTCATCCTTTTGGGTGCCAGAGACAAAGGAAG-AACTCTCGAAGAGAT 1657  
  
QY 534 ThrIlePheValSerLeuSerIle 541  
||||: |||  
Db 1658 ACAATGGTCTCTCCGCTGAGCCTT 1681

RESULT 2  
AK065497  
LOCUS  
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J013023P09, full insert sequence.  
AK065497 2266 bp mRNA linear PLN 24-JUL-2003  
ACCESSION  
VERSION AK065497.1 GI:32975515  
KEYWORDS FLI\_CDNA; CAP trapper.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE  
AUTHORS  
1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Doi,K., Kishimoto,N., Yazaki,J., Itoh,M., Hotta,I., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Sugiyama,A., Mizuno,K., Yokomizo,S., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,

Ikedo,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN: Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
Science 301 (5631), 376-379 (2003)  
2752273  
12869764  
2 (bases 1 to 2266)  
AUTHORS  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kurosaki,T., Kusumegi,T., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Matsuyama,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Murata,M., Nagata,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Nilkura,J., Nishi,K., Nomura,M., Namiki,T., Narikawa,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.  
Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
This clone is one of the 28K full-length cDNA clones from japonica rice.  
URL : http://cdna01.dna.affrc.go.jp/cDNA/  
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.  
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Nilkura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., yasunishi,A. and Hayashizaki,Y.  
Location/Qualifiers  
1. .2266  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol type="mRNA"  
/cultivar="Nipponbare"



/db\_xref="taxon:39947"  
/clone="J013023P09"

ORIGIN

Alignment Scores:

Pred. No.: 1.77e-141 Length: 2266  
Score: 2017.50 Matches: 427  
Percent Similarity: 70.07% Conservative: 46  
Best Local Similarity: 63.26% Mismatches: 52  
Query Match: 72.60% Indels: 151  
DB: 8 Gaps: 6

US-10-051-909-36 (1-553) x AK065497 (1-2266)

QY	3	SerSerSerPheArgProAlaGlyLysLysLysLysLysLysLysAsnGlnGlyLeuArg	22
Db	5	TCTCTCTCATCCATTTCTCCGGTTGAGCGCGCGGC-----	40
QY	23	ArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGly	42
Db	41	-----GGCGAGCCGAGCGGAGCGAGCGGCGCGCGCGATGAACGGCGCGCGG	88
QY	43	SerAsnArgGlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAsp-----	59
Db	89	AGCAGCGGGGAGGTCGTGTCGCCCGACGAGCGGCGACCGACCGCGCGGCGAGC	148
QY	60	---GlyValLeuArgArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArg	78
Db	149	GGCGCGGCTTGCGGAAGCCGCTGCTGAACACGGGGAGCTGGTACAGGATGGGATCG	208
QY	79	GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer	98
Db	209	TCCAGCCTCGCCGCC-----TCCTCCATGGCCGCCCATCCGGAGTCCCACGTC	259
QY	99	AlaPheLeuCysThrIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly	118
Db	260	GCCTTCCTCTGCACGCTCATCGTCGCGCTCGGCCCATCCAATTCGATTCCACCG	319
QY	119	PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe	138
Db	320	TTCTCCTCACCCACCGAGGACGCCATCATCCGCGACCTCAAGCTCTCCATCTCCG	379
QY	139	SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln	158
Db	380	TCGGCTTCGGTTCGCTGTCCAAACGTCGGCGCCATGGTCGGAGCGATCGCCAG	439
QY	159	MetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIle	178
Db	440	ATGGCGGAGTACATTGGCCGGAAGGGTCGTGATAATTGCGCGCGTTCCTAACAT	499
QY	179	GlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeu	198
Db	500	GGTTGGCTTGCCATCTCCTTTGCAAAAGACGCGTCATTTTATACATGGGACGCT	559
QY	199	GluGlyPheGlyValGlyIleIleSerTyrThr-----	209
Db	560	GAAGGTTTGGTGTGGTGTATATCATATACGGTCTGTATTACCAACTTCATTTCAT	619
QY	209	-----	209
Db	620	TTTATTGCTGCTTAGAATGTTCTGTGGCTTCACTGCTTGAGTACTTATCTGACTG	679
QY	209	-----	209
Db	680	TGAGATGCCTGTAGAATGAGGCAGAGCATCCTTTGCTAATGAAGCTTTTACAT	739
QY	209	-----	209
Db	740	TTGCCAGAAGAACTTTGGATCAGTTAAATAGTTAGTGCATGTATAAAACCTTAA	799
QY	210	-----ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGly	224
Db	800	ATTATATCGTACCAGGTGCCAGTATACATAGCAGAGATATCTCATCAGAAACAC	859

QY	225	AlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeu	244
Db	860	GCACTTGGCTCCGTGAACCAAGTTGTCCGTTACCATTTGGTATCTTGTGGCCCTATT	919
QY	245	GlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeu	264
Db	920	GGCATGTTTGTCTCTGGAGGCTGCTTGCACTGATAGGAAGCATCCCATGTACATTG	979
QY	265	IleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThr	284
Db	980	ATACCTGGTCTATTCTTCATTCGGAATCCCCAAGATGGCTGGCAAAATGAAATGAT	1039
QY	285	GluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGlu	304
Db	1040	GATGATTTTGAGGCTTCTCTACAAGTTTGGGGGGTTTGAGACTGATATTACTGCAG	1099
QY	305	ValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGln	324
Db	1100	GTGAATGATATAAAGAGAGAGCATCGGCAACAAAAAGAACACAGTCCGTTTAAA	1159
QY	325	GluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuLeuValLeu	344
Db	1160	GAGTTGAACCAAAAGAAATACCGCACTCCCTTACTGATAGGAACCTGGCCTTCT	1219
QY	345	GlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAla	364
Db	1220	CAGAACTAAAGTGAATAAATGATATCTGTTTATGCAAGTAGAATCTTCAGAGAT	1279
QY	365	Gly-----	365
Db	1280	GGTAAGTTGTGGATATCTCTTATAGGTCTCTTAGTACTACAGTCACGTTGTAGTAA	1339
QY	366	-----Val-ThrAsnSerAspLeuAlaThrCy	374
Db	1340	GGCTTGTAAATCTTAACACGCAAGAACTTCAGGGTTTACAAACAGTACTTGGCCAC	1399
QY	374	sSerLeuGlyAlaIleGln-----	380
Db	1400	TGCACCTAGGAGCAATTTCAGGTAAGTATTAGGAATATCCTGATCTCATCTCTGTAA	1459
QY	381	-----ValLeuAlaThrGlyValThrTh	388
Db	1460	TATCAAGTTGCTTACAGAAATCGTTGGATTACATTAGGTTCTTGGCACTGGAGTTACA	1519
QY	388	rTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetTh	408
Db	1520	ATGGTTACTGGACAGAGCTGGTCGACGGATGCTCCTTATTATCTCTACTGTGGGATG	1579
QY	408	rLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSe	428
Db	1580	TCATAAGCCTTCTTGCAGTATCTGTTGTATTTTCTGGAGGGTAATAATTTCCACATG	1639
QY	428	rAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleTh	448
Db	1640	TCATTTCGTTCTACATCTTAAAGTATGATCTCCTTGGTTGCTCTTGTGGCTTATATCAT	1699
QY	448	rPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSe	468
Db	1700	CTTTTCTCTCGGCATGGGTGCCATTCATGGGTCTGATGTCTGAGATCTCCAGTTAG	1759
QY	468	rIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIl	488
Db	1760	CATCAAGAGTCTTGGGGGAAGCTTTCGACACTTGCCAAACATGCTTACATCTCTGGGCA	1819
QY	488	eThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMe	508
Db	1820	AACAATGACAGCAAAATTTGTACTCAGCTGGAGTGTGGAGGAACATTTTGTCTCTAC	1879
QY	508	tValValSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly**	528
Db	1880	GATTGTGAGTGCTTTTCACTCTCGTGTTCGTATATTTTGGGTGCCAGAGACAAAAGGA	1939

Qy	528	*AsnSerArgGlyAspThrIlePheValSerLeuSerIleGln	542	
Db	1940	-AACTCTGGAGGAGATACAGTTTTCGTTCTGCTAAATTTGAG	1981	
RESULT 3				
AK068594				
LOCUS	AK068594	Oryza sativa (japonica cultivar-group)	cdna clone:J013154G15	full
DEFINITION		insert sequence.		
ACCESSION	AK068594			
VERSION	AK068594.1	GI:32978611		
KEYWORDS	FLI_CDNA; CAP trapper.			
SOURCE	Oryza sativa (japonica cultivar-group)			
ORGANISM	Oryza sativa (japonica cultivar-group)			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.			
REFERENCE	1			
AUTHORS	The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team; Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group:, Otomo,Y., Murakami,K., Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y., Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M., Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J., Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J., Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:, Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y., Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Yoshino,M. and Hayashizaki,Y.			
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice			
JOURNAL	Science	301 (5631),	376-379	(2003)
MEDLINE	22752273			
PUBMED	12869764			
REFERENCE	2	(bases 1 to 1948)		
AUTHORS	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.			
TITLE	Direct Submission			
JOURNAL	Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)			
COMMENT	This clone is one of the 28K full-length cDNA clones from japonica rice. URL : http://cdna01.dna.affrc.go.jp/cdna/ NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,			

Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.				
FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J., Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S., Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.				
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.				
FEATURES	Location/Qualifiers			
source	1..1948			
	/organism="Oryza sativa (japonica cultivar-group)"			
	/mol_type="mRNA"			
	/cultivar="Nipponbare"			
	/db_xref="taxon:39947"			
	/clone="J013154G15"			
ORIGIN				
Alignment Scores:				
Pred. No.:	6.95e-127	Length:	1948	
Score:	1821.50	Matches:	366	
Percent Similarity:	79.93%	Conservative:	72	
Best Local Similarity:	66.79%	Mismatches:	99	
Query Match:	65.55%	Indels:	12	
DB:	8	Gaps:	4	
US-10-051-909-36 (1-553) x AK068594 (1-1948)				
Qy	1	ProSerSerSerSerPheArgProAlaGlyLysLysLysLysLysLysAsnGlnGly	20	
Db	17	CCCCTCAATCAATCAAGCTTAAGCTAAGCTGGCGGCACAGCGAGAGAAACGAGGGAT	76	
Qy	21	LeuArgArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGly	40	
Db	77	CGATCGCGTCCACGCACGCCCGAT---CCCGCGCGAGATGAGTTCGCGGCGGAGGAG	133	
Qy	41	GlyGlySerAsnArgGlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAspGly	60	
Db	134	AGCGGCGCGAGGATGGGGGCGGCGGCGG-----TCGGCGTCCGAC-----	175	
Qy	61	ValLeuArgArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSer	80	
Db	176	---CTGCGGAAGCCGTTCTCTCCACACGGGGAGCTGGTACAAGATGTCGTGCGGCGCGC	232	
Qy	81	SerPheAlaProGlyThr-----SerSerMetAlaValLeuArgGluSerHisVal	97	
Db	233	GGCGGCGGGATGGGGTTCGGCTCGGCTCCTCCGCTACTCCCTCCGCGACTCCTCCGTC	292	
Qy	98	SerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSer	117	
Db	293	TCCGCCGTCCTCTGCACCCCTCATCGTCGCCCTCGGCCCCCATCCAGTTCGGCTTCACCTGC	352	
Qy	118	GlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu	137	
Db	353	GGCCTCTCTCGCCACCCAGGACGCCATCATCTCCGACCTCGGCCTCACCTCTCCGAG	412	
Qy	138	PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGly	157	
Db	413	TTCTCGCTGTTGGGTGCTGTTCGAACGTCGGGGCAATGGTGGGCGCCATCGCCAGCGGC	472	





Db 229 ---GGAGATCTGCGGAAACCAATTTTTCGATACAGGAAGTTGGTATAGAAATGGGTTCCAGA 285

Qy 79 GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98

Db 286 CAATCAAGTTTG-----ATGGGTTCTTCTCAAGTTATTTCGAGAAAGCTCCATCTCT 336

Qy 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118

Db 337 GTCCTTGCTTGTTCTTATTATTGTTGCTCTTGCTCTATTTCAGTTTGGTTTACTGCTGGT 396

Qy 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138

Db 397 TATTCATCTCCAACCTCAATCAGCAATTACAAGTGAACCTTGGTCTTTCTGTGCAGAGTAC 456

Qy 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158

Db 457 TCATGGTTTGGTTCCTCTATCCAATGTGGTGCCATGGTTGGGCAATTGCAAGTGGTCAA 516

Qy 159 MetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIle 178

Db 517 ATATCTGAGTATATTGGAAGAAAGGTCCTTTGATGATTGCTGCTATCCCCAACATAATT 576

Qy 179 GlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeu 198

Db 577 GGATGGCTTGCTATATCATTTGCCAAAGATCTTTCATTTCTGTATATGGGAAGGATGTTG 636

Qy 199 GluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSer 218

Db 637 GAAGGTTTGGCGTGGGAATAATATCTTACACGGTTCCTGTGTACATTTCTGAAATAGCA 696

Qy 219 ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238

Db 697 CCACAGAATTGAGGGGAGCTTTGGGTTTCAGTAAACCAAGCTATCCGTTACAATCGGTATA 756

Qy 239 PheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAla 258

Db 757 ATGCTATCGTATATGCTGGGCTTTTGTCCGTGGAGGATACCTTGCACTTTTAGGAATT 816

Qy 259 LeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeu 278

Db 817 TTACCATGCACTATATTGATACCTGGCTTATCTTCATACCGGAATCTCCTAGTGGTTG 876

Qy 279 AlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu 298

Db 877 GCGAAAATGGGTATGATGGAAGAAATTGAGGTTTCCTTGCAAGTTCTCCGGGGCTTTGAC 936

Qy 299 ThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArg 318

Db 937 ACTGATATATCTCTTGAAGTTAATGAATCAAGAGATCTGTTGCTTCAAGTAGCAAAAGA 996

Qy 319 ThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGly 338

Db 997 ACCACCATAGAATTGCAGAACTACGACAAAGAAGATATTGGTTACCTTGATGATAGGG 1056

Qy 339 IleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSer 358

Db 1057 AATGGTTTACTTATTTTGCAACAACCTCAGTGAATCAATGGTGTCTCTTCTATTCTAGT 1116

Qy 359 SerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAla 378

Db 1117 ACCATCTTTAAAGAGGCTGGGTCCACATCAAGCAACGCTGCAACATTTGGCCTTGGTGCT 1176

Qy 379 IleGlnValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIle 398

Db 1177 GTTCAGGTTATCGCAACTGTTGTGACTACCTGGTTGGTGGATAAATCAGGACGCAGACTT 1236

Qy 399 LeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhe 418

Db 1237 CTCTCTAATTGCTCTTCTCTGGGATGACATGAGTCTTCTAGTGGTCGCAATGTCCTTC 1296

Qy 419 PheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSer 438

Db 1297 TTTTAAAGGAATGGTATCAGACGAATCTACATGGTACAGTGTATTCTCCATCCTTTCA 1356

Qy 439 LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp 458

Db 1357 GTTGTGGAGTTGGCTATGGTGGTACTTTCTCACTAGGAATGGAGCAATTCGTTGG 1416

Qy 459 LeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThr 478

Db 1417 ATTATAATGTGAGAGATTCTTCCAATTAAACATCAAGGGGCTGGCTGGAAGTATTGCTACA 1476

Qy 479 LeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrp 498

Db 1477 CTTGCCAACTGGTTTGTAGCATGGATTGTGACAATGACTGCAAAACATAAATGTTGTATGG 1536

Qy 499 SerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheVal 518

Db 1537 AATAGTGGAGGTACTTTCTCTATTATACATGGTTGTTGTGCTTTCACTGTGGCCTTTGTG 1596

Qy 519 ValLeuTrpValProGluThrLysGly\*\*\*AsnSerArgGlyAspThrIlePheValSer 538

Db 1597 GTCATTTGGTCCCCGAGACAAGGGGAG-AACATTTGGAGGAGATTCAGTG---GTCCTT 1652

Qy 539 LeuSerIleGlnArgGlnLeuGlnTrpLeuProGluCys 551

Db 1653 CAGAAGATAAAGAAATTTTCTGATAACTGTCCATGGTGT 1691

RESULT 5

AY093274 1557 bp mRNA linear PLN 21-APR-2002

LOCUS Arabidopsis thaliana similar to integral membrane protein

DEFINITION (At1g19450) mRNA, complete cds.

ACCESSION AY093274

VERSION AY093274.1 GI:20259851

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1557)

Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.

Direct Submission

TITLE Submitted (26-MAR-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

JOURNAL

COMMENT e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Southwick,A., Nguyen,M., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K., Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P., Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

Southwick,A., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

location/Qualifiers

1. .1557

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="1"

source









QY	212	ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln	231
Db	492	GTATATATCGCTGAGATAGCTCCACAAATATGAGAGGAGATTGGGTTTCGGTTAACCCAG	551
QY	232	LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArg	251
Db	552	CTTTCTGTGACAATTGGAATAATGTTGGCGTATTACTTTGGCCTCTTTGTTCATGGAGA	611
QY	252	LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle	271
Db	612	ATTCTTGCAAGTTCTGGGGATATTGCCGTGTACATTATTGATACCAGGTCCTTTTTCATC	671
QY	272	ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu	291
Db	672	CCCGAATCCCCCTCGCTGGCTGGCAAAATGGGTATGACAGATGAATTTGAAACTTCATTA	731
QY	292	GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla	311
Db	732	CAAGTTCCTCGAGGATTTGAGACTGATATTACCGTTGAAGTTAATGAATCAAGAGATCT	791
QY	312	ValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr	331
Db	792	GTGGCATCATCTACAAAGCGAAATACAGTTCGGTTTGATATCTCAACCCAGGAGATAC	851
QY	332	ArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn	351
Db	852	TATTTCCCACTTATGGTTGGTATAGGGTTGCTGTACTTCAACAACCTTGGTGGAAATTAAT	911
QY	352	GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu	371
Db	912	GGTGTCTCTATTCAGTACATAATTTGAATCTGCAGGAGTTACATCGAGTAACGCA	971
QY	372	AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeu	391
Db	972	GCAACATTTGGSGTCGGGGCTATTCAGGTAGTAGCGACTGCAATATCCACTTGGTTGGTG	1031
QY	392	AspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu	411
Db	1032	GACAAAGCAGGTGCTGGCTTCTGCTTACTATCTCTTCGGTTGGGATGACGATTAGCCTT	1091
QY	412	LeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr	431
Db	1092	GTAATTGTTGCAGTGCTTTCTATCTTAAGGAATTTGTGTCTCCTGTGATTCAGACATGTAC	1151
QY	432	TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe	451
Db	1152	AGTTGGCTGAGCATATTGTCCAGTAGTTGGAGTGTGGCAATGGTTGTCTTTTCTCATTG	1211
QY	452	GlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSer	471
Db	1212	GGAATGGGACCAATACCGTGGCTCATATTATGTCTGAGATCCTTCCTGTGAACATAAAGGT	1271
QY	472	LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr	491
Db	1272	TTAGCTGGAAGTATTGCAACTCTAGCCAATTGGTTCTTTTCTTGGTTGATCACCATGACA	1331
QY	492	ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSer	511
Db	1332	GCAATTTGCTGTAGCCTGGAGCAGTGGAGAACTTTCACCTCTGTATGGATTGTTTGT	1391
QY	512	AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArg	531
Db	1392	GCATTACAGTGGTGTTCGTGACTCTATGGGTTCCTGAGACCAAGGCAA-AACTCTTGA	1450
QY	532	GlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeuGlnTrpLeuProGluCys	551
Db	1451	AGAACTTCAATCCTTGTTCAGATGAACAATTTGAACAACACTTCATCTTGTCACCCTCT	1510
QY	552	LeuSer	553
Db	1511	CTCTCC	1516

AY124845	LOCUS	AY124845	1464 bp	mRNA	linear	PLN 06-JUL-2002
Arabidopsis thaliana	DEFINITION	Arabidopsis thaliana	At1g75220/F22H5_6	mRNA	complete cds.	
AY124845	ACCESSION	AY124845				
AY124845.1	VERSION	AY124845.1	GI:21700860			
FLI CDNA.	KEYWORDS	FLI CDNA.				
Arabidopsis thaliana (thale cress)	SOURCE	Arabidopsis thaliana (thale cress)				
Arabidopsis thaliana	ORGANISM	Arabidopsis thaliana				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.						
1 (bases 1 to 1464)	REFERENCE					
Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.	AUTHORS					
Arabidopsis ORF clones						
Unpublished						
2 (bases 1 to 1464)	REFERENCE					
Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.	AUTHORS					
Direct Submission						
Submitted (17-JUN-2002)	TITLE					
SIGnAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	JOURNAL					
The two base pair differences do not change the amino acid sequence of the protein.						
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.	COMMENT					
The Salk, Stanford, PGEN (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Shinn,P., Banh,J., Bowser,L., Chan,M.M., Chang,E., Dale,J.M., Deng, J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura, Y., Yu,G., Davis,R.W., Theologis,A., and Ecker,J.R.						
Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.						
location/Qualifiers	FEATURES					
1. .1464	source					
/organism="Arabidopsis thaliana"						
/mol_type="mRNA"						
/db_xref="taxon:3702"						
/chromosome="1"						
/clone="U13497"						
/ecotype="Columbia"						
/note="This clone is in pUNI 51"						
1. .1464	CDS					
/note="putative integral membrane protein"						
/codon_start=1						
/product="At1g75220/F22H5_6"						
/protein_id="AAW70554.1"						
/db_xref="GI:21700861"						
/translation="MSFRDNEEARNDLRRPFHTGSWYRMGSRQSSMMGSSQVIRDS SISVLACVLIVALGPIQFGFTCGYSSPTQAAITKDLGLTVSEYSVFGSLSNVGMVGA						







QY 511 rAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly\*\*\*AsnSerAr 531  
Db 1633 TGGATTCACTGTGGTTTGTGAGTCTTTGGGTCTCTGAGACTAAGGAAA-AACGCTTG 1691  
QY 531 gGlyAspThrIlePheValSerLeu 539  
Db 1692 AAGAGATCCAAGCTTTGTTCAGATG 1716  
RESULT 10  
AX653119 1377 bp DNA linear PAT 22-MAR-2003  
LOCUS Sequence 2989 from Patent WO03000898.  
DEFINITION AX653119  
ACCESSION AX653119  
VERSION AX653119.1 GI:29155933  
KEYWORDS Oryza sativa  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1  
AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,  
Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 2989 03-JAN-2003;  
Syngenta Participations AG (CH)  
FEATURES  
source  
1. .1377  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.71e-119 Length: 1377  
Score: 1710.00 Matches: 358  
Percent Similarity: 78.36% Conservative: 44  
Best Local Similarity: 69.79% Mismatches: 39  
Query Match: 61.53% Indels: 73  
DB: 6 Gaps: 9  
US-10-051-909-36 (1-553) x AX653119 (1-1377)  
QY 44 AsnArg---GlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAsp-----Gly 60  
Db 4 AACAGGGCGCGCGCGTCGCGCTCGCGGATGAGAGCGGAAGCGACTACGAGAGCGGCGGC 63  
QY 61 ValLeuArgArgProLeuLeu---AsnThrGlySerTrpTrpArgMetSerSerArgGln 79  
Db 64 GGCATGCGGAAGCCGCTGCTGATGCACACGGGGAGCTGGTACAGGATGGGTCGCGGCAG 123  
QY 80 SerSerPhe---AlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98  
Db 124 GGGAGCCTCACCGGCGCGGGACCTCGTCCATGGCCATCTACGCGAGTCCCACGTCTCC 183  
QY 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118  
Db 184 GCCTTCCTCTGCACGCTCATCGTCGCGCTCGGCCCCCATCCAGTTCGGAATCACGGGGGC 243  
QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138  
Db 244 TTCTCCTCCCGACGCGAGGACCCATCATCCGAGACCTCGACCTCACCTCTCCGAGTTC 303  
QY 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158  
Db 304 TCGGTGTTTCGGATCGCTGTCCAACGTCGGCGCCATGGTTGGGGCGATGCCAGTGTCAA 363  
QY 159 MetAlaGluTyrIleGlyArgLysGly----- 167  
Db 364 ATGCCCGAGTACATTGGGCGCAAAGGGKGTGAGAAATTTTTCCTTTTGGGTACACATC 423  
QY 168 SerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys 187

Db 424 TCATTGATGATTGCTGCAATTCCAAACATCATTTGGTTGGCTTGCCATCTCCTTTGCAGAAA 483  
QY 188 AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSer 207  
Db 484 GACTCATCGTTTCTTTATATGGGACGATTGCTCGAGGGGTTGGTGTGTGTCATCTCT 543  
QY 208 TyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGly 227  
Db 544 TATACGGTGCCAGTTTACATAGCAGAAATATCACCTCAAAACATGAGAGGIGCTCTTGGC 603  
QY 228 SerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe 247  
Db 604 TCAGTGAATCAGTTATCTGTAAACCGTTGGTATATTGTGGCATATTGTCTCGSCATGTTT 663  
QY 248 IleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267  
Db 664 GTTCCTTGAGAGGCTTCTTGCTGTAATAGGAATCTTGCCTTGCACTGTGTTGATACCTGGC 723  
QY 268 LeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCys 287  
Db 724 CTATTCTTCATTCCAGAAATCCCAAGATGGTTGGCAAAAGATGAACATGATGATGATTTT 783  
QY 288 GluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAsp 307  
Db 784 GAGACTTCTTTACAAGTTCTGAGGGGATTTGAGACTGACATCAGCGCGGAAGTGAATGAT 843  
QY 308 IleLysArgAlaValAlaSerSerLysArgThrThrIleSer---PheGlnGluLeu 326  
Db 844 ATAAAG-----GTAGCATCTTCAAAGCAGCAGTCATATTTTAGCCGATTTGGCTTCGTG 897  
QY 327 AsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsn 346  
Db 898 ACACGCAAA----- 906  
QY 347 LeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyVal 366  
Db 907 -----ATTTCTTCAGGTCTC 921  
QY 367 ThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyVal 386  
Db 922 ACAAACAGTGACTTGGCTACATGTGCACCTTGGTGTATCCAG----- 963  
QY 387 ThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGly 406  
Db 964 -----ATCTCTCTCTGCTGGG 978  
QY 407 MetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGln 426  
Db 979 ATGACTCTAAGCCTCCTTGCAGTTGCTGTTGATTTTCTCAAGGATAGCATTCACAA 1038  
QY 427 AspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheVal 446  
Db 1039 GATTCTCAGATGTACTACACCTTAAAGTATGATCTCCTTGGTTGCTCTTGTGGCTTTGTA 1098  
QY 447 IleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuPro 466  
Db 1099 ATCGCCTTCTCTTCGGTATGGGTGCCATTCCATGGATCATATAATGTGAGATCCTCCCG 1158  
QY 467 ValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPhe 486  
Db 1159 GTTAGTATCAAGAGTCTCGCAGGAAGCTTTGCGACGCTCGCCAACTGGCTTACATCCTTT 1218  
QY 487 AlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSer 506  
Db 1219 GGAATAACAATGACAGCAAACTTGATGCTTAGCTGGAGTGTGGAGGGACCTTTGTGTCC 1278  
QY 507 TyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThrLys 526  
Db 1279 TACATGGTCGTGAGTGTCTTACCCTCGTGTTCGTCTATCTCTTGGGTGCCAGAGACAAA 1338  
QY 527 Gly\*\*\*AsnSerArgGlyAspThrIlePheValSerLeu 539  
Db 1339 GGAAG-AACTCTCGAAGAGAGATACAATGGTCTCTCCGCTG 1376

```

RESULT 11
BT014234
LOCUS      1825 bp      mRNA      linear      PLN 11-MAY-2004
DEFINITION Lycopersicon esculentum clone 133433R, mRNA sequence.
ACCESSION  BT014234
VERSION     BT014234.1  GI:47105649
KEYWORDS    FLI_CDNA.
SOURCE      Lycopersicon esculentum (tomato)
ORGANISM    Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1 (bases 1 to 1825)
AUTHORS     Kirkness,E.F., Wang,W. and Vazeille,A.
TITLE       Direct Submission
JOURNAL     Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
            Medical Center Drive, Rockville, MD 20850, USA

FEATURES             Location/Qualifiers
     source            1..1825
                        /organism="Lycopersicon esculentum"
                        /mol_type="mRNA"
                        /db_xref="taxon:4081"
                        /clone="133433R"
                        /tissue type="fruit"
                        /note="TZXHM68"

ORIGIN
Alignment Scores:
Pred. No.:      2,7e-117      Length:      1825
Score:          1693.00      Matches:     328
Percent Similarity:  82.85%      Conservative:  73
Best local Similarity: 67.77%      Mismatches:   79
Query Match:      60.92%      Indels:       5
DB:              8              Gaps:         2

US-10-051-909-36 (1-553) x BT014234 (1-1825)

Qy   57  AspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySerTrpTyrArgMetSer 76
      ::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   122 AACAAATGAAGGGGATCTGAAAAGCCATTTTACATACTGGGAGTTGGTATCGTATGGGG 181

Qy   77  SerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHis 96
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   182 TCAAGACAGTCTAGTATG-----ATGGGTCTCTCTCAAGCTATTCTGTGATAGTTCT 232

Qy   97  ValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThr 116
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   233 ATTTCTGTTCTTGCTGTGTCTCATAGTTGCTTTGGGCCCAATCCAGTTTGGTTTCACT 292

Qy   117 SerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSer 136
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   293 TGTGGCTACACTTACCAACGCAAAATTGCAATTACCAAAAGATCTGAAGCTCACTGTTTCT 352

Qy   137 GluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSer 156
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   353 GAGTTCTCTTTGTTGGTTCCTTTATCCAATGTGGGTGCTATGGTTGGTGAATATCTAGT 412

Qy   157 GlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsn 176
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   413 GGTCAAAATTTCTGAGTACATTGGCGAAAAGGGTCTTTAATGATAGTGTCTATACCAAAAT 472

Qy   177 IleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArg 196
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   473 ATCATTTGTTGGCTTGCATCTCATTTGCCAAAGATCCTTCTTTCTATACATGGGAAGA 532

Qy   197 LeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGlu 216
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   533 TTGTTGGAGGGCTTTGGTGTGGGAATAATCTCTTATGTGGTTCCTGTATATATGCTGAG 592

Qy   217 IleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPhe 236
      ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   593 ATTGCACCTCAGATAATGAGAGGGGGTCTAGGGTCTGTAAACCAAGCTCTCTGTACAATT 652

```



ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
1  
REFERENCE  
AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,  
Katagiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 2990 03-JAN-2003;  
Syngenta Participations AG (CH)  
FEATURES  
source Location/Qualifiers  
1..1428  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.43e-109 Length: 1428  
Score: 1577.00 Matches: 318  
Percent Similarity: 85.85% Conservative: 34  
Best Local Similarity: 77.56% Mismatches: 32  
Query Match: 56.75% Indels: 27  
DB: 6 Gaps: 2  
US-10-051-909-36 (1-553) x AX653120 (1-1428)  
QY 130 AspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGly 149  
Db 277 GATCTGTACTTACAACAGCATATGTTCTCGGCTTTCGGTTCGCTGTCCAACGTCGGCGCC 336  
QY 150 MetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeu 169  
Db 337 ATGGTCGGAGCGGATGCCAGTGGCGAGATGGCGAGTACATTTGGCCGGAAGGGTCGTTG 396  
QY 170 MetIleAlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAla 189  
Db 397 ATAATTGGCGCGGTTCCTAACATCATTTGGTTGGCTTGCCATCTCCTTTGCCAAAAGACGCG 456  
QY 190 SerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209  
Db 457 TCATTTTATACATGGGACGCTTGCTTGAAGGTTTGGTGTTGGTGTATATACATATACG 516  
QY 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229  
Db 517 GTGCCAGTATACATAGCAGAGATATCTCATCAGAACACACAGAGGAGCACTTGGCTCCGTG 576  
QY 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIlePro 249  
Db 577 AACCAGTTGCCGTACCATTGGTATCTTGTGGCCTATTGCTAGGCATGTTGTCTCT 636  
QY 250 TrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhe 269  
Db 637 TGGAGGCTGCTTGCAGTGATAGGAAGCATCCCATGTACATTTGTTAATACCTGGTCTATT 696  
QY 270 PheIleProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThr 289  
Db 697 TTCATTCCGGAATCCCCAAGATGGCTGGCAAAAATGAAATGATGGATGATTTTGAGGCT 756  
QY 290 SerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLys 309  
Db 757 TCTCTACAAGTTTGGGGGGTTTGAGACTGATATTACTGCAGAA----- 801  
QY 310 ArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLys 329  
Db 802 AGAGCAGTAGCATCGGCAACAAAGAACCCACAGTCCGTTTAAAGAGTTGAACCAAAAG 861  
QY 330 LysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGly 349  
Db 862 AAATACCGCACTCCCTTACTGATAGAACTGGCCTTCTGTACTTTCAGAAATCTAAGTGA 921  
QY 350 IleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSer 369

Db 922 ATAAATCGTATTCTGTTTTATGCAAGTAGAATCTTTCAGAGATGCAGGGTTTACAAACAGT 981  
QY 370 AspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTyr 389  
Db 982 GACTTGGCCACATGTGCACTAGGAGCAATTTCAG----- 1014  
QY 390 LeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeu 409  
Db 1015 -----ATCTCTACTGCTGGGATGACTCTA 1038  
QY 410 CysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsn 429  
Db 1039 AGCCTTCTTGCAGTATCTGTTGTATTTTCTGGAGGGTAATATTTCACATGATTCTCAT 1098  
QY 430 SerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPhe 449  
Db 1099 TCGTTCTACATCTTAAGTATGATCTCCTTGGTTGCTCTTGTGGCTTATATCATCACCTTT 1158  
QY 450 SerPheGlyMetGlyAlaIleProTyrLeuMetMetSerGluIleLeuProValSerIle 469  
Db 1159 TCCTTCGGCATGGTGGCATTCATTCATGGTTCATGATGCTGAGATCCTCCAGTTAGCATC 1218  
QY 470 LysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTyrLeuThrSerPheAlaIleThr 489  
Db 1219 AAGAGTCTTGGGGGAAGCTTTCGACACATTTGCCAACATGCTTACATCCTGGCAATAACA 1278  
QY 490 MetThrThrAsnLeuMetLeuThrTyrSerValGlyGlyThrPheLeuSerTyrMetVal 509  
Db 1279 ATGACAGCAAAATTTTACTCAGTTCAGTGGAGTGGAGGAACATTTTGTCTCTACATGATT 1338  
QY 510 ValSerAlaPheThrIleValPheValValLeuTyrValProGluThrLysGly\*\*\*Asn 529  
Db 1339 GTGAGTGTCTTCACTCTCGTGTTCGTCAATATTTTGGGTGCCAGAGACAAAAGGAAG-AAC 1397  
QY 530 SerArgGlyAspThrIlePheValSerLeu 539  
Db 1398 TCTGGAGGAGATACAGTTTTCGTTCTGCTA 1427  
RESULT 13  
AC120986 204649 bp DNA linear PLN 02-SEP-2004  
LOCUS Oryza sativa (japonica cultivar-group) chromosome 5 clone  
DEFINITION OJ1781 H11, complete sequence.  
ACCESSION AC120986  
VERSION AC120986.2 GI:46518554  
KEYWORDS HTG.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
REFERENCE 1 (bases 1 to 204649)  
AUTHORS Chow,T.-Y., Hsing,Y.-I.C., Chen,C.-S., Chen,H.-H., Liu,S.-M.,  
Chao,Y.-T., Chang,S.-J., Chen,H.-C., Chen,S.-K., Chen,T.-R.,  
Chen,Y.-L., Cheng,C.-H., Chung,C.-I., Han,S.-Y., Hsiao,S.-H.,  
Hsiung,J.-N., Hsu,C.-H., Huang,J.-J., Kau,P.-I., Lee,M.-C.,  
Leu,H.-L., Li,Y.-F., Lin,S.-J., Lin,Y.-C., Wu,S.-W., Yu,C.-Y.,  
Yu,S.-W., Wu,H.-P. and Shaw,J.-F.  
TITLE Oryza sativa BAC OJ1781\_H11 genomic sequence  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 204649)  
AUTHORS Chow,T.-Y. and Hsing,Y.-I.C.  
TITLE Direct Submission  
JOURNAL Submitted (16-MAY-2002) Institute of Botany, Academia Sinica, 128,  
Section 2, Academia Road, Nankang, Taipei 11529, Taiwan  
REFERENCE 3 (bases 1 to 204649)  
AUTHORS Hsing,Y.-I.C. and Chow,T.-Y.  
TITLE Direct Submission  
JOURNAL Submitted (29-MAR-2003) Institute of Botany, Academia Sinica, 128,  
Section 2, Yien-chu-yuan Road, Nankang, Taipei 11529, Taiwan  
REFERENCE 4 (bases 1 to 204649)  
AUTHORS Chow,T.-Y.  
TITLE Direct Submission



DKLGRIADWTRNPAGPAALAAASRDSVDFDFTSVDDSLAAAEADSSFRLVDKPPPRHP  
RFGPKWFENQRPQLPORRDEEVEARRREAEKERARRERHFQNNRSHHHHPGRGNQSSS  
AKSSVDIQDWTMRREQIPFANFTKLSFSVADQPEDLLCGAVEYYDRAFDVRNPKAAR  
RLERFKSRNFFKVTITDDPVIIRLAEEDKATVFATDAILAALMCTPRSISHSWDIVVQR  
VGNKLFDDKRDGSQLDLLSVNETAQQLPENKDDINSASHSLAVEATYINQNFSSQOVL  
LDGEKVTDFENPFASEGEEAASGVYRRWKLDDEISIVARCEVHAVNADPGGRQF  
LTLNALNEFDPKITGVDRQKLETORGAVLATELTKNNANKLARWTQALLAGADMML  
GYSVSRPHDRHNHAILTVMGYKPRDFATQINLTSNMWGIKSVIDI CMKPEEGKYV  
LVKDPAPKPOVRIYEVPSDAFENDYVVEEPLPEEEQVRPPSDVDVDATEEMDAAEAEAN  
NAAASAGGEKSAAAAA"  
complement (65704. .66784)  
/gene="OJ1781\_H11.7"  
/note="unknown protein"  
complement (join(65704. .65774,65857. .65914,66247. .66306,  
66414. .66461,66620. .66682,66764. .66784))  
/gene="OJ1781\_H11.7"  
/note="similar to rice EST AK058475"  
/codon\_start=1  
/product="unknown protein"  
/protein\_id="AAU10659.1"  
/db\_xref="GI:51854278"

Alignment Scores:

Pred. No.: 1.68e-96 Length: 204649  
Score: 1453.50 Matches: 426  
Percent Similarity: 37.12% Conservative: 45  
Best Local Similarity: 33.57% Mismatches: 54  
Query Match: 52.30% Indels: 747  
DB: 8 Gaps: 19

US-10-051-909-36 (1-553) x AC120986 (1-204649)

Qy	3	SerSerSerSerPheArgProAlaGlyLysLysLysLysLysAsnGlnGlyLeuArg	22
Db	118274	TCTCTCTCATCCATTCTCCGGTTGAGCGGCGCCGC	118309
Qy	23	ArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGly	42
Db	118310	-----GGCGGAGCGGAGCGGAGCGAGCGGCGGCGGCGGATGAACGCGCGCGG	118357
Qy	43	SerAsnArgGlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAsp	59
Db	118358	AGCAGGCGGGAGGTGCGTCTGCGCCGACGAGCGGCGGACGACGAGCGGCGGAGC	118417
Qy	60	---GlyValLeuArgArgProLeuLeuAsnThrGlySerTyrArgMetSerSerArg	78
Db	118418	GGCGGCGGCTTGCGGAAGCCGCTGCTGAACACGGGGAGCTGTGTACAGGATCGCGG	118477
Qy	79	GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer	98
Db	118478	TCCAGCCTCGCCGCC-----TCCTCCATGGCCCGCCATCCGGGAGTCCCACGTCCTC	118528
Qy	99	AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly	118
Db	118529	GCCTTCCTCTGCACGCTCATCGTCGGCTCGGCCCCCATCCCAATCGGATTCACCA	118588
Qy	119	PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu	137
Db	118589	TTCTCCTCACCCACCGAGCGCCATCATCCGCGACCTCAAGCTCTCCATCTCCGAGGTC	118648
Qy	137	-----	137
Db	118649	CCGTGCCCGTTCAGCTGCGATCTTTCGCTGCTTTTTCGCTTCCTCTTGCTTAATGCGTT	118708
Qy	138	-----PheSerAlaPheGlySerLeuSerAs	146
Db	118709	TCTTGTAGTGTGAAGTTTATCTTTCTGATGCAGTTTCTCGGCTTTCGGTTCGCTGTCCA	118768
Qy	146	nValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgGly	166
Db	118769	CGTCGGCGCCATGGTCGGAGCGATCGCCAGTGGGCGAGATGGCGGAGTACATTGGCGCGAA	118828
Qy	166	sGly-----	167

Db	118829	AGGGGTGAGAATTTGTGCTTCATTTGGTCCCTTATCTGTCTTCCACTAGAAGGCCATCCT	118888
Qy	167	-----	167
Db	118889	GTATCTCGACGACGAACAATTGATACTTTGATAGTTTGACAAAGGTCAATACGTCACTGTTG	118948
Qy	167	-----	167
Db	118949	TATGATTTCCGGTAACACCAATAAATTTGCCATTAAAGAAAGTGAGTGGGATTAGAAATGC	119008
Qy	167	-----	167
Db	119009	TGGTCATTTCTGTGAACCTGFAGCTTGTGAAAAATATGAAACCCCAAAATTTTCAGGTTCA	119068
Qy	167	-----	167
Db	119069	TAGCTATACTATTGGGGTTTGACCTTAATTTCAAAGTTGACGAGTTCCTTTAGTTGTGATT	119128
Qy	167	-----	167
Db	119129	GCCATTGATCTGGCAAAAGAGGAGTTAGTATGACACAGAAAATGTGTAGATGAAGTTTGT	119188
Qy	167	-----	167
Db	119189	GACATTTTCTTTGGTTCACTTCATAATCACTTTAATCAGATCAGTAGTTAACTACAGTG	119248
Qy	168	-----SerLeuMetIleAl	172
Db	119249	AAGCTCTATTATTTTACTCTAATGGTCTTGCTGTGATTGAATCTCCAGTCGTTGATAATGC	119308
Qy	172	aAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys	187
Db	119309	GGCGGTTCTTAACATCATTCGTTGGCTTGCCATCTCCTTTGCAAAAGTACCGTGTACATA	119368
Qy	187	-----	187
Db	119369	TTTCTCAAGCTTATTTTCTCTGTGTTGCTCTTTTCACTAGTTTCACTTTCGCCGATCTT	119428
Qy	188	-----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValG	204
Db	119429	TGTTTTTCAGGACGCGTCATTTTATACATGGGACGCTTGCTTGAAGGTTTGGTGTGG	119488
Qy	204	YIleIleSerTyrThr	209
Db	119489	TGTTATATCATATACGGTCTGTATTTACCAACTTCATTTATGTTTATTTGCTGCTTAGA	119548
Qy	209	-----	209
Db	119549	ATGTTCTGTGGCTTCACCTGTGAGTACTTATCTGACTGAGACTGAGATGCCTGTAGAAT	119608
Qy	209	-----	209
Db	119609	GAGGACAGCATCCTTTGCCTAAATGAAGCTTTTACATTTTGTGTCAGAAAGAACTTTG	119668
Qy	210	-----Va	210
Db	119669	GATCAGTTAATAGTTAGTCATGTATAAACCCCTTAATGTGTGAATTATATCGTACCAGGT	119728
Qy	210	lProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValas	230
Db	119729	GCCAGTATACATAGCAGAGATATCTCATCAGAAACACAAGAGGACACTTGGCTCCGTGAA	119788
Qy	230	nGln-----	231
Db	119789	CCAGGTGCTTATCGTCAGAAAGCATCTTATCTTATCGAGGAAATTTATGTTTGT	119848
Qy	232	-----LeuSerValThrPheGlyIlePheL	240
Db	119849	ATGTTGGACTTACCATGCAGTTTTTTCATTATGTAGTTGTCCGTTACCATTTGGTATCTTGT	119908
Qy	240	euAlaTyrLeuLeuGlyMetPheIleProTrrpArgLeuLeuAlaValIle	256



Db 119909 TGGCCATATTTGCTAGGCATGTTTGTTCCTTGGAGGCTGCTTGCAGTGATAGGTACAAATC 119968  
QY 256 ----- 256  
Db 119969 AATACCTGATGCTGAATGTTTCTATTAGGGATATGTAATAATAATAATGAATTTT 120028  
QY 256 ----- 256  
Db 120029 AACTTTTGATATTGCTAGTACCTTTTGGACTTATTGCAATTTTGATGTGCGTTAAATGT 120088  
QY 257 -----GlyAlaLeuProC 261  
Db 120089 CTGGTACCTCTGATTCTGATTCAAATATCTGGGTTGCCATGCTTGTAGGAAGCATCCCAT 120148  
QY 261 YsThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeu----- 278  
Db 120149 GTACATTGTTAATACCTGGTCTATTCTTCATTCCGGAATCCCCAAGATGGCTGGTATGTT 120208  
QY 278 ----- 278  
Db 120209 ATACACCTAAGTTGGATATTGTATCACAAATCTGAAATCTGTAGCAAGCACTTTCCTACA 120268  
QY 279 -----AlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuA 295  
Db 120269 TCTATATGCAGGCAGCAAAATGAAAATGATGGATGATTTTGAGGCTTCTCTACAAGTTTGA 120328  
QY 295 xgGlyPheGluThrAspIleThrThrGluValAsnAspIleLys----- 309  
Db 120329 GGGGGTTTGAGACTGATATTACTGCAGAAAGTAATGATATATAAAGGCAAGCAATTTTATTT 120388  
QY 309 ----- 309  
Db 120389 TTCTTGATTAACTGACTACAAGTTATTTCGGGTTTCGGGTGATTTTCATCTTAGTAAGATT 120448  
QY 309 ----- 309  
Db 120449 TTCACTAACACAAACACTTGGGAGATATGTACTTTTCATTTTCTGGAAGACTAATATGTT 120508  
QY 310 -----ArgAlaValAlaSerSerSerLysArgThrThrIleSerPhe 323  
Db 120509 CAAATCCCTGTTATGCAGAGAGCAGTAGCATCGGCAACAAAGAACACACAGTCCGTTTT 120568  
QY 324 GlnGluLeuAsnGlnLysTyrArgThrProLeuLeuLeu----- 337  
Db 120569 AARGAGTTGAACCAAAAGAAATACCGCACTCCCTTACTGGT - AACTTACATACCCCTTTT 120627  
QY 337 ----- 337  
Db 120628 GTTCAGTTTATACCAAAAGTCATTAATCATGTTTGTAAATGGTAACAGCTTTATGTTGA 120687  
QY 338 -----GlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe 355  
Db 120688 CAGATAGGAACGTGCCTTCTTGTACTTCAGAATCTAAGTGAATAAATGGTATTCTGTTT 120747  
QY 356 TyrAlaSerSerIlePheLysAlaAlaGly----- 365  
Db 120748 TATGCAAGTAGAATCTTCAGAGATGCAGGTAAGTTGTGGATATCTCTTATAGGTCCTTAG 120807  
QY 365 ----- 365  
Db 120808 TACTACAGTCACGTTGTAGTAAATTTTGGCTTGTAATCTTAAACACGCAAGAACTTTCAGG 120867  
QY 366 Val-ThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln----- 380  
Db 120868 GTTTACAAACAGTGACTTGGCCACATGTGCCATAGTAGGACAAATTCAGGTAAGTGATTAGGA 120927  
QY 380 ----- 380  
Db 120928 ATATCCTGATCTCATCTCTGTAACTTATATCAAGTTGCTTACAGAAATCGTTGGATTACA 120987  
QY 381 -----ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLe 399  
Db 120988 TTAGGTTCTTGCCACTGGAGGTTACAACATGGTTACTGGACAGAGCTGGTCGACGGATGCT 121047

QY 399 uLeuIle----- 401  
Db 121048 CCTTATTGTAATTTCTCTTCCCAATGAAATCACCTTTTTTGTAGTGTCTGTAACAATA 121107  
QY 402 -----IleSerThrSerGlyMetThrLeuCysLeuLeuA 413  
Db 121108 ACATAGTGACTCACTTAAAAATTGCAGATCTCTACTGCTGGGATGACTCTAAGCCTTCTTG 121167  
QY 413 laValSerValValPhePheLeu----- 420  
Db 121168 CAGTATCTGTTGTATTTTCTGGAGGTACATTTCAGATTTTTCAGTATACCATTGTGA 121227  
QY 420 ----- 420  
Db 121228 TAAAAAAAATCATATGTTGCAGTCTCTTTGATAGCCTACATTGTCGCATGGACTCTTCCC 121287  
QY 421 -----LysAspA 423  
Db 121288 CTTTCTCCTCAGTTCACCTCAATATTTCAATATTACTACAATGATTATTTGTTTCAGGGTA 121347  
QY 423 snIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleV 443  
Db 121348 ATATTTACATGATTTCTCATTCGTTCTACATCTTAAAGTATGATCTCCTTGGTTGCTCTTG 121407  
QY 443 al----- 443  
Db 121408 TGGTATGCTGTGAGTTAAACCAATTTCTTTTGAAAAATTGCTCCAAAAACCTCTGCCGGAAA 121467  
QY 444 -----SerPheValIleThrPheSerPheGlyMetGlyAla 455  
Db 121468 ATTTACCTCGGTCACTTACTTCAGGCTTATATCATCACCTTTTCTTCCGCATGGGTGCC 121527  
QY 456 IleProTrpLeuMetMetSerGluIle----- 464  
Db 121528 ATTCATGGGTCATGATGCTGAGGT - ATAATGCTTTTGCCCCCATAGCTCACAATGATGTT 121586  
QY 464 ----- 464  
Db 121587 AGTATTTTGTAGAGCATTTGACACATCCACCATTTTTGTGCAAGTATCTCAGGCATCAGCTT 121646  
QY 465 -----LeuProValSerIle 469  
Db 121647 AAGTAATCTAAAGATTTTCAGTTAAATCTTTGTTTCTTAGATCCTCCAGTTAGCATC 121706  
QY 470 LysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThr 489  
Db 121707 AAGAGTCTTGGGGGAAGCTTTTGGCACACTTGCCAAACATGCTTACATCCTGGGCAATAACA 121766  
QY 490 MetThrThrAsnLeuMetLeuThrTrpSerValGly----- 501  
Db 121767 ATGACAGCAAAATTTGTTACTCAGCTGGAGTGCTGGAGGTCCTGATTTCTATAAAAATCCT 121826  
QY 501 ----- 501  
Db 121827 CTGTTACCAAGAATAAAAATTCGAACTGATTTACCTGATAAAAAGTTAAATATTTTATTTT 121886  
QY 502 -----GlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheVa 518  
Db 121887 AATTTCAACAGGAACATTTTGTCTCTACATGATTTGTAGTGCTTTCACCTCTCGTGTTCGT 121946  
QY 518 lValLeuTrpValProGluThrLysGly\*\*\*AsnSerArgGlyAspThrIlePheValSe 538  
Db 121947 CATATTTTGGGTGCCAGAGACAAAGGGAAG - AACTCTGGAGGAGATACAGTTTTCGTTCT 122005  
QY 538 rLeuSerIleGln 542  
Db 122006 GCTAAATTTTGAG 122018

RESULT 14

AC025808/c

LOCUS

DEFINITION

Genomic sequence for Arabidopsis thaliana BAC F18014 from  
120977 bp DNA linear PLN 11-OCT-2000

ACCESSION VERSION KEYWORDS SOURCE ORGANISM	chromosome 1, complete sequence. AC025808 AC025808.8 GI:7636235 HTG. Arabidopsis thaliana (thale cress) Arabidopsis thaliana Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE AUTHORS	1 (bases 1 to 120977) Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,Q., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A., Theologis,A. and Ecker,J.R. Genomic sequence for Arabidopsis thaliana BAC F18O14 from chromosome 1
TITLE	Unpublished
JOURNAL	2 (bases 1 to 120977)
REFERENCE AUTHORS	Ecker,J.R.
TITLE	Direct Submission
JOURNAL	Submitted (15-MAR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE AUTHORS	3 (bases 1 to 120977) Ecker,J.R.
TITLE	Direct Submission
JOURNAL	Submitted (22-APR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE AUTHORS	4 (bases 1 to 120977) Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Nguyen,M., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
TITLE	Direct Submission
JOURNAL	Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
REFERENCE AUTHORS	5 (bases 1 to 120977) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shinn,P., Altafi,H., Bei,B., Chin,C., Chiou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
TITLE	Direct Submission
JOURNAL	Submitted (11-OCT-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
COMMENT FEATURES	On Apr 22, 2000 this sequence version replaced gi:7543634. Location/Qualifiers 1..120977 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /db_xref="taxon:3702" /chromosome="1" /clone="F18O14" 90..2159 /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F18O14.1" /protein_id="AAF79419.1" /db_xref="GI:8778411" /translation="MISFEVSPDVFTCSIVVNAYCRSGNVDDKAMVFAKETESSLGLELNVVTYNSLINGYAMIGDVEGTRVRLMSERGVSRNVVYTTSLIKGYCKKGLMEEAEH
source	
CDS	

VFELLKEKLVADQHMYGVLMDGYCRTGQIRDVAVRVDHNMIEIGVVRTNTTICNSLING YCKSGQLVEAEQIFSRMNDWSLKPDDHHTYNTLVDGYCRAGYVDEALKLCDQMCQKEVV PTWMTYNILLKGYSRIGAFHDVLSLWKMKLKGVDNAEISCSSTLLEALFKLQDFNEAM KLWENVLARGLLTDTITLNVMSIGLCKMEKNEAKEILDNVNIFRCKPAVQTYQALSH GYKVGNLKEAFVKEYMERKGIPTTIEMYNLTILISGAFYKHLNKVADLVIELRARGL TPTVATYGALITGCNIGMIDKAYATCFEMIEKGITLNVNICSKIANSFLRLDKIDEA CULLQKIVDFDGLLPGYQSKFELEASATCLTKIQIAESVENPCKLLVPNNIVYN VAIAGLCKAGKLEDARKLFSDLLSSDRFIPDEYTYTILIHGCAIAGDINKAFTLRDEM ALKGIIPNIVTYNALIKGLCKLGNVDRAQLLHLKLPQKGITPNAITYNTLIDGLVKSG NVAEAMRLKEKMIKGLVRGSDKQGDVDIPKEVVLDPEVKLGSTGVIEMNNSNELYDVR RVSEAVV"	
complement (4647..5702) /note="putative glycosyltransferase-like protein emb CAB42905.1; similar to EST gb AI998490.1" /codon_start=1 /evidence=not_experimental /product="F18O14.2" /protein_id="AAF79456.1" /db_xref="GI:8778448"	CDS
/translation="MSQHLLLLILLSLLLHKKPISATTIIQKFKEAPQFYNsADCPLI DSESDDDVVAKPIFCsRRAVHVAMTLDAAVIRGSVAAVLsVLQHSsCPENIVeHFVA SASADASSLRATISSsPPYLDFTVYVFNVSsVSRLISSsIRsALDCPLNYSYsLADL LPPCVRrVVYLDsLILVDIDIAKLAATDLGRDSVLAAPeYCNANFTsYFTsFWSNPt LSLTfADRKAcyFNTGVMVIDLSrWRREGAYTSRIeEWMAMQKMRiYELGSLLPpFLLV FAGLIKpVNRWNOHGLGGDNfRGLCRDLHPGPVSLLLHWSGKGKpWARLdAGRPcPLD ALWAPYDLLQTPeALDS"	
complement (9620..10300) /note="similar to Zn finger protein gi 6319911" /codon_start=1 /evidence=not_experimental /product="F18O14.3" /protein_id="AAF79455.1" /db_xref="GI:8778447"	CDS
/translation="MSDVPSCSSGNDTNNNDSSNFECNICLDLAQDPiVTLcGHLfCW PCLYkWLHLHSQSKDCpVCkAVIEEDRLVPLYGRGKSSADPRKSIPGLeVNPSPsGQ RPETAQPPDPNHGFAHHHGfGGMGGFAAPMASARfGNVTLSAAGGLIPSLfNLHfH GFPDAAMyGAASGGfPHGfSNPFHGCHSHMSYQRHTGRGQGDHHLRIILLIIVfVV VVFSLfLS"	
join(12523..12577,12686..13374) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F18O14.4" /protein_id="AAF79420.1" /db_xref="GI:8778412"	CDS
/translation="MAIFSTSHLLFIATCTISVSGTTFITLHNcGStIWPGILt ANGAQLGDGGfALASGSsVTFtVSPGWSGRFWARTYCNfDfASGSCKGTGDcGSKLKC AGAGGAPPATfLAFTIGSSGKKNAVQDPYDVsLVdGYNVQMKITPQGGSGDCKTAGCV SDVNAICPKeLOVtGPSVAAcKsACAFNKPEYcCTGAYsTPATcPTNYSKIFKQA CPSAYSAYDDASSTfTCTNANYEISfCS" complement (join(14020..14136,14236..14304,14402..14464, 14557..14696,14879..15005,16274..16486)) /note="unknown protein; similar to EST gb H76493.1" /codon_start=1 /evidence=not_experimental /product="F18O14.5" /protein_id="AAF79454.1" /db_xref="GI:8778446"	
/translation="MLEAVDSSGVVNGGfPQIOsFYGDcSSEELsVLPRHTKVVVtG NNRTKsVLVGLOGVKKAVGLGGWHVLVTNGIEVKLQRNALsVLEPPTGNEEDDdLD FENTORNGSDMTsEDTLKPHKsKLRQSSRSsSHKTMRSLSsDSQSsSGTTPPENM KVDLSKLEMPALLNYWRHfNLVDAlPNfPSKEQLIDIVQRHfMSQQMDelQVIvGFVQA AKRMKACkFQsKESRNTDLNCIS" join(17994..18489,18568..18698,18814..18847,19006..19099, 19170..19347,19441..19479,19616..19715,19855..19994) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F18O14.6" /protein_id="AAF79421.1" /db_xref="GI:8778413"	CDS
/translation="MAKTDKLAQFLDSGIYESDEFNWFFLDTVRITNRSYTRfKVsPS AYSRFFNSKQLNQHSSesNPkKRKKQKNsSFHLPSVGEQASNLRHQEARLFLSKAH	

ESFLKBIELLSLTKGLSDDDDSSLLNKKCCDDDEVSFIELGGVWQAPFYEITLSFNL  
VCNNEGESCNEQRFVFNVLVVEIGEEVEAEFNSCFYMSDLHHRNL  
HPCNHEGYNLIVIDPPWENASAHQSKFVIYLMYPTLPNQYFLSLPIKQLAHABGAL  
VALWYTNREKLLSFVEKELFPAGIKYVATMYLWKVPDGTLICDLLLAGSEKXSDP  
KLLDKNQIIMSIPGDFSRKPPIGDILLKHTPGSQPARCLELFAREMAAGWTSWGNPL  
HFQDSRYFLKV"

CDS

join(22037. .22300,22618. .23361)  
/note="unknown protein; similar to ESTs gb|AI992723.1,  
gb|T41777.1, dbj|AV440980.1, gb|AI997341.1, and  
dbj|AV440785.1"  
/codon\_start=1  
/evidence=not\_experimental  
/product="F18O14.7"  
/protein\_id="AAF79422.1"  
/db\_xref="GI:8778414"  
/translation="MTSDGATSTSAAAAAAMATRRKPSWRENNRRRRRAVAA  
KIYTLRAQGNYNLPKHCDDNEVLKALCSEAGWVVEEDGTYRKGHKPLPGDMAGSSS  
RATPYSSHNQSPLSSTFDSPILSYQVSPSSSPSPSRVGDPHNISTIFPFLRNGGIP  
SSLPLRISNSAPVTPPVSSPTS RNPKPLPTWESFTKQSMMAAQSMTSLNYPFYAV  
SAPASPTHHRQFHAPATIPCEDSSSTVDSGHWISFOKPAQQPPFSASMVPTSPTFN  
LVKPAPQQLSPNTAAIQEIGQSSEFKFENSQVKPWEGERIHDVAMEDLELTIGNGKAH  
S"

CDS

complement(join(23868. .24983,25237. .25407))  
/note="unknown protein; similar to EST gb|T13858.1,  
emb|Z38050.1, and emb|Z38087.1"  
/codon\_start=1  
/evidence=not\_experimental  
/product="F18O14.8"  
/protein\_id="AAF79453.1"  
/db\_xref="GI:8778445"  
/translation="MAGRRDRSQQLRGSRIAAILIGIPGVCVAVLPYPGFNSSSS  
LKASEHLSKSSNOVGSSACESPERVMLKSDFTLSEKNAELKKQVRELTEKRLAEQ  
GSDNARKQVLALGTQIKAGPFTVKSRLTNPTILPDESINPRLAKILEEIAVDKEVIV  
ALANANVAMLEVQIASIKRVGITYNLVVALDDYIENLCKENDVAYYKRPDKDVTIV  
GKTGGNHAVSGLKFRVLRFLQLGYGVLLSDVDIVFLQNPFSHLYRDSDVESMSDGH  
NHTAYGFNDVDFDEPAMGWARYAHTMRIWVNSGFFYLRPTIPSIELLDRAVADRLSKAK  
VMDQAVFNEELFYPSHPETHALHASKRVMDMYEFMNSKVLFKTVRKNHELKKVKPVI  
VHNYHPDKLNRMQAVVEFYNGKQDALDSFPDGS"

CDS

complement(26100. .27917)  
/note="unknown protein; similar to ESTs gb|T42717.1,  
emb|F13886.1, and emb|F13885.1"  
/codon\_start=1  
/evidence=not\_experimental  
/product="F18O14.9"  
/protein\_id="AAF79452.1"  
/db\_xref="GI:8778444"  
/translation="MADPLNGKSPFICFSLLSFTLLFISPLYATESPVIEDVSTDVA  
VSVSETNREAVLLHNLEELVKNLTVELVANLDAKLSATPLKEKNEISVDDDDIGEEXRG  
RAKAPSTVKYSPFWSERFQSAVKLNSDATCINLVPFDPFEGSSKYFAIGDSKGRVY  
VFLRNGDVLIEFFTVDSPVTAMVSYSSVFKNSSFFVTGHQNGAVLLHRIHEGNGED  
WNSNSVSMEHVGFVDVDSADPVTLLVHVHVRVRYILATDLSGKLTVLTENRTVYGS  
VIPSSRPLVFLKQRLFLFTESGAGSLDLRSMKIRETECEGLNHSLARTYVFDAAERSK

Alignment Scores:

Pred. No.:	3.85e-73	Length:	120977
Score:	1137.50	Matches:	335
Percent Similarity:	40.58%	Conservative:	68
Best Local Similarity:	33.74%	Mismatches:	78
Query Match:	40.93%	Indels:	518
DB:	8	Gaps:	18

US-10-051-909-36 (1-553) x AC025808 (1-120977)

QY	52	GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer	71
Db	67805	GAAGAGGGAAGAAATGAT-----CTTCGGCGACCGTTCTTACACACAGGAAGT	67758
QY	72	TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
Db	67757	TGGTATCGGATGGGTCGAGACAATCTAGTATG-----TTGGAATCGTCTCAAGTT	67707
QY	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProile	111
		:::    :::    :::	

Db	67706	ATTCGAGACAGCTCAATTTCTGTCTTAGCTTGTGTTTTGATTGTTGCTCTTGGTCCTATT	67647
QY	112	GlnPheGlyPheThrSer-----	117
Db	67646	CAATTCGGATTCACTGT-AAGAACAACATAAAACAATAATATCTCAACTTCTTATATTTTT	67588
QY	117	-----	117
Db	67587	TTTATCAATGTGGATTTCTCTGCTTTTGTCTCAAAATTGACTAAGAAAGTGTGATTTTGAT	67528
QY	118	-----GlyPheSerSerProThrGlnAspAlaMetValArgAsp	130
Db	67527	TTATCTTTTGTTCAGTGTGGTTATTCTCACCACACTCAAGCTGCCATTACTAAGGAT	67468
QY	131	LeuAenLeuSerIleSerGlu-----	137
Db	67467	CTTGGTTTAACTGTATCAGAGGTATGAAGGTCCCATCTTTATTTAAGAACTGTTCTGTTT	67408
QY	137	-----	137
Db	67407	TTTTTTTTTTTTTTGTAAAGACTCTTGTCTGTAAATTTTCTGACTTTTAGGTTTCTTGT	67348
QY	138	-----PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAla	153
Db	67347	TTGATGTTTTTAGTACTCTGTGTTTGGGTCTCTATCCAAATGTGGGTGCTATGGTTGGAGCA	67288
QY	154	IleAlaSerGlyGlnMetAlaGluTyrIleGlyArg-LysGly~	167
Db	67287	ATTGCGAGTGGTCAGATTGCTGAATACGTTGGACGGAAGGGGTATAATAGTTTGTGTTTC	67228
QY	167	-----	167
Db	67227	TTGATTTCTAAACATGTTCTCATCTGTGACAAATTTGTCATGTAAACATCTTCTTCTTGTT	67168
QY	168	-----SerLeuMetIleAlaAlaIlePr	175
Db	67167	GTAATCTTATTACTGTGTTTTTCTTCTTCTTTTGACAGTCTCTGTATGATTGCTGCAATTCC	67108
QY	175	oAsnIleIleGlyTyrLeuAlaIleSerPheAla-Lys-----	187
Db	67107	CAATATTATTGGATGGCTTTCGATATCATTTGCAAAAGTGAGTGTGTTGGTCTCTCAAA	67048
QY	187	-----	187
Db	67047	GTTTCTGATATTCTTGGCTTATCTCTATACCTACTAAGTTGTTGTTGTTCTTGGTTTTTG	66988
QY	188	-----AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleI	206
Db	66987	TTTAGGATACTTCTTTTCTTACATGGGAAGATTGTTAGAAGGCTTCGGAGTTGGGATTA	66928
QY	206	IeSerTyrThrVal-----	210
Db	66927	TCCTCTATACGGT-AATATTAGATTCAACTATTTCTTTTGATTTCAGCAATGTTGTTTGG	66869
QY	211	-----ProValTyrIleAlaG	216
Db	66868	AGTGATCAATATAATTACACACCTTTTGAATTGCGATTTTAGGTTCCCGTATATATAGCTG	66809
QY	216	IuIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAenGlnLeu-----	232
Db	66808	AGATCGCTCCACAGACCATGAGAGGAGCCTTAGGTTTACGTTAACACAGGT-TTGTCATGTT	66750
QY	232	-----	232
Db	66749	TAGTTAGCTTCACTATTTTATGTTTTCATTTTATGAGTCGATTCATATGTTAAACACTAA	66690
QY	233	-----SerValThrPheGlyIlePheLeuAlaTyrLeuL	244
Db	66689	CAATAAATGTTTGTGTTTTATAGCTTCTGTGAACAATTGGGATAAATGCTGGCGTATTTAC	66630
QY	244	euGlyMetPheIleProTyrArgLeuLeuAlaVal-----	255
Db	66629	TCGGTCTCTTTGTTCCATGGAGAATTCTTGCAGTTTTCGGTAACACTCACAAAACATTTTGA	66570



QY 255 ----- 255  
Db 66569 CTTGACTTGGGAGAACTATACCTGTTTATTATTATTTTGTATCGATACAGTCTAATCAC 66510  
QY 256 -----IleGlyAlaLeuProCysThrMetLeuIleProGlyL 268  
Db 66509 TTTCTGTGTGCTCTCTTATGTTTAGGAGTATGCGCATGTACATTGTTGATACCGGGTC 66450  
QY 268 euPhePheIleProGluSerProArgTrpLeu----- 278  
Db 66449 TTTTYYTCATTCCUGAATCTCCTCGGTGGTTGGTATGCCTCTACAACCGTATTGCTTATA 66390  
QY 279 -----Ala 279  
Db 66389 TTGTTGAGTTTGTGTCATTATTGTTATCCTAAATTTGTTGGTGTGTAATTGTAGGCA 66330  
QY 280 LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr 299  
Db 66329 AAGATGGGTTTGACAGATGATTTCGAAACTTCATGCAAGTCTTCGTGGATTGAGACT 66270  
QY 300 AspIleThrThrGluValAsnAspIleLys----- 309  
Db 66269 GATATTACCGTTGAGGTTAATGAATCAAGGTAGTGACTAAACTGAAGAAATGCTTTTGAT 66210  
QY 310 -----ArgAlaValAlaSe 314  
Db 66209 GTGATTTCTTGCTTATTAAAAATTTTCATTTTATGCACTCTGCAGAGATCTGTGGCATC 66150  
QY 314 rSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrPr 334  
Db 66149 ATCTAGCAAAACGTTCTGCAGTTCGGTTTGTAGACCTCAAGCGCAGGAGATACTATTCCC 66090  
QY 334 oLeuLeu----- 336  
Db 66089 ACTGATGGTGCAATTTCTCTTACCTCTCTTGAACATAATTTTGTCAACACATGAGAGAAAC 66030  
QY 337 -----LeuGlyIleGlyLeuLeuValLeuGlnAsnL 347  
Db 66029 TCACTCTTTTCCTGTAATCTGTGTTACAGGTGGTATAGGGCTGCTTGCACTTCAACAAC 65970  
QY 347 euSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAa----- 364  
Db 65969 TTGGAGGAATCAATCGTGTCTTGTCTATTTCGAGTACAATTTTGAATCTGC-AGGTAGA 65911  
QY 364 ----- 364  
Db 65910 GTCATATGTTTTTTTGTGATTATTGGAATGGTCTTAAGATAGATAAAACAGATTTCCTAACT 65851  
QY 365 -----GlyValThrAsnSerAspLeuAlaThrC 374  
Db 65850 ATGCTTCTTTGTTGCAATTATCATTCACCTAAGGGGTGACATCAAGTAATGTGGCGACAT 65791  
QY 374 ysSerLeuGlyAlaIleGlnValLeu----- 382  
Db 65790 TTGGAGTTGGCGGTTGTTTCAGGTATA-TCATTTCAATCTAGGAGACATTAGATTCGATTTT 65732  
QY 383 -----AlaThrGlyValThrThrTrpLeuLeuA 392  
Db 65731 TCTAACCGTTGTTTCATTCTATACCAGGTAGTGGCGACTGGGATAGCGACATGGTTGGTGG 65672  
QY 392 spArgAlaGlyArgArgIleLeuLeulle----- 401  
Db 65671 ATAAAGCAGGTCGTCGGCTTCTACTCATGCTGGAAGAATATTTGAGATTTCATCTGTATATC 65612  
QY 401 ----- 401  
Db 65611 TTACAAAGTCTGAAGAATTTATGTATGCAATAACATAATATTAGCATTTTCTTCTGAAATG 65552  
QY 402 ---IleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeu 420  
Db 65551 CAGATCTCTTCTATCGGAATGACGATAAGCCTCGTGTATTGTAGCAGTTGCATTTTACCTT 65492

QY 421 Lys----- 421  
Db 65491 AAGGTAAACATTTTAACTTTTGACATCTCTACTTGTACAATGTAAGAGTGTCACTCTAATT 65432  
QY 422 -----AspAsnIleSerGlnAspSerAsnSer 430  
Db 65431 ATGAGATATTGAGGTTTCTTTAATCTTTTGTTAGGAATTTGTATCACCTGATTCCAACATG 65372  
QY 431 TyrTyrIleLeuThrMetIleSerLeuValGlyIleVal----- 443  
Db 65371 TACAACATTTCTAAGCATGGTTTCCGTAGTTGGAGTTGTGGTAAGTCGTTTGTTTATACA 65312  
QY 443 ----- 443  
Db 65311 TTCACACAACGAAAGATACTTCTTTATTATTACATATTTCCCGGTATGTGTACTAAATCTGG 65252  
QY 444 -----SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp 458  
Db 65251 CTGAGTTATTACAGGCTATGTTATTCTTGTCTCTTAGGAATGGGACCAATTCATGG 65192  
QY 459 LeuMetMetSerGlu----- 463  
Db 65191 CTGATTATGTCTGAGGTACTAAATATTAAAAACATGGTGTTCCTTAGTGTAAAAATCACAA 65132  
QY 464 -----IleLeuProValSerIleLysSe 471  
Db 65131 CTGGAATTTTAATACTACTGATTAAATGTTGGTAAACAGATTTCTGCCAGTGAATATAAAGG 65072  
QY 471 rLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetTh 491  
Db 65071 TTTAGCCGGAAGTATAGCAACTTGTTTAAACTGGTTCGTGTCTCATGGTTAGTGACAAATGAC 65012  
QY 491 rThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSe 511  
Db 65011 TGCAATATGCTGTTTAGCATGGAGCAGTGGGGGAACTTTCACCTCTCTACGCTTTGGTTT 64952  
QY 511 rAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly\*\*\*AsnSerAr 531  
Db 64951 TGGATTCACTGTGGTTTGTGAGTCTTTGGGTTCTCGAGACTAAAGGAAA-AACGCTTG 64893  
QY 531 gGlyAspThrIlePheValSerLeu 539  
Db 64892 AAGAGATCCAAGCTTTGTTCCAGATG 64868  
RESULT 15  
AC025814 68041 bp DNA linear PLN 22-JAN-2001  
LOCUS Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence,  
DEFINITION complete sequence.  
ACCESSION AC025814  
VERSION AC025814.7 GI:12331602  
KEYWORDS HTG.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 68041)  
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M.-I., Creasy,T.H., Haas,B.J.,  
Wu,D., Maiti,R., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T.R.,  
Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.  
Arabidopsis thaliana chromosome 1 BAC F22H5 genomic sequence  
Unpublished  
REFERENCE 2 (bases 1 to 68041)  
AUTHORS Lin,X. and Kaul,S.  
TITLE Direct Submission  
JOURNAL Submitted (15-MAR-2000) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, xlin@tigr.org  
REFERENCE 3 (bases 1 to 68041)  
AUTHORS Town,C.D. and Kaul,S.  
TITLE Direct Submission  
JOURNAL Submitted (12-SEP-2000) The Institute for Genomic Research, 9712  
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org

REFERENCE AUTHORS TITLE JOURNAL	4 (bases 1 to 68041) Town,C.D. and Kaul,S. Direct Submission Submitted (05-OCT-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
REFERENCE AUTHORS TITLE JOURNAL	5 (bases 1 to 68041) Town,C.D. and Kaul,S. Direct Submission Submitted (12-OCT-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
REFERENCE AUTHORS TITLE JOURNAL COMMENT	6 (bases 1 to 68041) Town,C.D. PhD. Direct Submission Submitted (22-JAN-2001) On Jan 22, 2001 this sequence version replaced gi:12280794. Address all correspondence to:at@tigr.org
	BAC clone F22H5 is from Arabidopsis thaliana chromosome 1 The orientation of the sequence is from SP6 to T7 end of the BAC clone. Genes were identified by a combination of several methods: Gene prediction programs including Genscan+ (Chris Burge, http://CCR-081.mit.edu/GENSCAN.html), GeneMarkHM (Mark Borodovsky, http://genemark.biology.gatech.edu/GeneMark/), GlimmerA (a variant of GlimmerM, see Mihaela Pertea, http://www.tigr.org/softlab/glimmerm.htm/glimmerm.html, and GeneSplicer (Mihaela Pertea and Steven Salzberg, contact mpertea@tigr.org), searches of the complete sequence against a peptide database and the plant EST database at TIGR (http://www.tigr.org/tdb/tgi.shtml). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/). Simple repeats are identified by repeatmasker (Arian Smit, http://ftp.genome.washington.edu/RM/RepeatMasker.html).
FEATURES source	1. .68041 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /cultivar="Columbia" /db_xref="taxon:3702" /Chromosome="1" /Clone="F22H5" complement(418. .5404) /gene="F22H5.15" complement(join(<418. .594,684. .773,873. .938,1034. .1095, 1179. .1227,1319. .4820,5040. .>5404)) /gene="F22H5.15" complement(join(<418. .594,684. .773,873. .938,1034. .1095, 1179. .1227,1319. .4820,5040. .5404)) /gene="F22H5.15" /codon_start=1 /product="unknown protein, 3' partial; 5120-134" /protein_id="AAG12690.1" /db_xref="GI:10092277" /translation="MEYEKSPATTTFPSRKISNNRSHLSFSANAVYDGVFSSPVNSKS PLVDYGEIFRGSGPSSIPFLDVPENLVGKVKVDVRSKLDYSSVFGLGACDPAVT PKVVIKSEKTSINEDKKRRKGGNSDVLPCNEGKSPFEMVRMKHSDISYHQTVP RNENGATHLTQVPATMPGPIPTQVVDNTSLHLKIESKSTPIPAVEKKLPCNEGREEVK ASRKQGSKTEVDENIFARDGSTRDDSTCKTVSNGEYRDVKPPSSFQCTLNGEHGAS ERLSGLNSGSPSERYETEDADSPSPYPFDAETDENSVAESSAALKKAIEEAQIRMNI AKQWMEKKKSGFRSCAKLKSCDSDKIENKGTKEGITEBSRDNQSIILGEMVKPSEQ SFSENGDQHAKRARKLWGPVEGLLKSTSDHKPEELEEQDIITLEEQARRGRKHHELP GGIFKSVMNSKQEPENLAPAKPEPDTKQEVQPIITENPEFTFGQLGSKLKCVEAFTG SKVYQKDEKQFTEKENSTVTQMVQDESDSQEMLAGIPIETYLREVEETPQQTFSKS EMNIQKSESTICAFTERSQQNMEKETGWQKSAKCFEDSGVKDFQENGDTCTNVD QEGSKDIVSEPQEMLVGPDSTKYVREVEETPTPLNKTQSDSDSVGAMVSFNRNISE PGNIDEVQEAHVKVPRRRRVWKTSEDVYNNMIKAPKGNRNPWQLESAENETTAMSFEE
gene	
mRNA	
CDS	
gene	GVRIHHASEEIESTGQASDSGLQENWTVLKQMFQMTADTKGEDETGYCLVESERG HLDIHQAEQKEQVEVETVRTNYEAYAHTRENEDESQAQETYCRKEDGKVEVQGKTSL VRELIGEELEMASLEEDVQEASEEAGWVQGLSEINEINEHADSHAEMLEYDRSETDS NNSRERFDQTQEQAEETMIDGSIIDTTSRSSFEMRQGDSYIEEVGIEQHRSDQPEKA SAVSNTEEHIEEIDSDSIQSGWSVVEDDDRSLODGASQAESKSHDELETKEEDEMK TSLGVERNGDKKELEHQFECQENETYSRVEAESSCRFPNGEEIIGAATNGNMKENE GEEGESCRSSMEEGDATSDIGAATDGNMKEQESCRSSMEEGDATSDISONKA ETVEEHLKKIDTETREKERERKQERVVERAIRERARAFADAMERAGKTAMEKAKAVA HRREVPKSEKSGSVENDKLSSAEKASMQAKLRARAARAVERAITEVRERAMEKALS GK SAASQAKSYGGSKSFSSGERRGSSSGTSENKSGPSSNQGTGEPIQRKARSERHQ RTSDRAAEALAEKKLRLDKTQKEQTERNLRAELADADVWRSSGKENNLRALISTIQY ILGAESGWKPIPLTDLVSSASVRKAYRKATLYVHPDKLQQRGASTQQKYICEKVPDLL K"
gene	complement(9354. .10899) /gene="F22H5.16" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<9354. .9638,9750. .9953,10043. .10328, 10461. .10595,10671. .>10899)) /gene="F22H5.16" complement(join(9481. .9638,9750. .9953,10043. .10328, 10461. .10595,10671. .10856)) /gene="F22H5.16" /codon_start=1 /product="NADPH oxidoreductase, putative; 10572-9197" /protein_id="AAG12695.1" /db_xref="GI:10092282" /translation="MTTEKSKILVIGGTGYMGFEIVEGSAKAGNPTFALVREASLSDP VKSKTIQSFKDLGVTLHGDLNDHESLVKAIKQVDVVISTIGHKQIFDQTKIISAIKE AGNVKRFLEPAEFGIDVERTSAVEPAKSLFAGKQIRRAIEAEGIPYTYVVSNCAGFY LRTLLQFESGLISHTRDKAIIFGDKNVPRDKVILGDGNKAVINKEEDVAAVMTKA VDDLRTLNKTLTYSPPNNILSMNEMVTLWEKKIGKSLKTHISEEQILKSIQVPIDVF KSINHAFVFKGQDQTSFTIEPWFGEESVLYPDVKYTSIDEYLSQFT" complement(11122. .12573) /gene="F22H5.18" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<11122. .11422,11516. .11734,11791. .12034, 12117. .12251,12333. .>12573)) /gene="F22H5.18" complement(join(11235. .11422,11516. .11734,11791. .12034, 12117. .12251,12333. .12518)) /gene="F22H5.18" /codon_start=1 /product="NADPH oxidoreductase, putative; 12234-10951" /protein_id="AAG12680.1" /db_xref="GI:10092267" /translation="MASEKSKILVIGGTGHIGKLIIEASVKAGHSTLALVREASLSDP NKGKTQNFKDFGVTLLHGDLNDHESLVKAIKQADVISTVGSQMILDQTKIISAIKE AGNVKRFLEPSEFGMDVDSASAVEPAKSAFGRKLQTRRDIEAEGIPYTYLVNYPAGY LPTLVQLEPGLTSPPRDKVKIFGDNVKVEYFTYAVINKEEDIAAYTIKAVDDPRTLNK TLYINPNNTLSMNEIVTLWEKKIGKSVKEIYMSEEQIFKSIQESVPFNVLLSINHA VFVKGDTNFTIEPSFGFEASELYPDIKYTSIDEYLSYFALGTSLNT" complement(12872. .14426) /gene="F22H5.17" /note="similar to GI:1708420 from [Arabidopsis thaliana] (J. Biol. Chem. 270 (44), 26224-26231 (1995))" complement(join(<12872. .13216,13298. .13513,13578. .13821, 13974. .14108,14193. .>14426)) /gene="F22H5.17" complement(join(13053. .13216,13298. .13513,13578. .13821, 13974. .14108,14193. .14378)) /gene="F22H5.17" /codon_start=1 /product="NADPH oxidoreductase, putative; 14094-12769" /protein_id="AAG12677.1" /db_xref="GI:10092264" /translation="MATEKSKILVIGGTGYIGKFLVEASAKAGHSTFALVREATLSDP VKGKTQSFKOLGVTLHGDLNDHESLVKAIKQVDVVISTVGSQMILDQTKIISAIKE AGNVKRFLEPSEFGVDVRTSAVEPAKSAFAGKIQIRRTIEAEGIPYTYAVDPPRTLNKI LPTLVQFEPLGTSPPRDKVTILGDGNAKEYFAAVINKEEDIAAYTIKAVDDPRTLNKI LYIKPSNNTLSMNEIVTLWEKKIGKSLKTHLPPEQLKSIQESPIPINVLNINHAV FVNGDTNISIEPSFGVEASELYPDVKYTSVDEYLSYFA"

```
gene      15081. 16153
/gene="F22H5.1"
/note="similar to GI:6939839 from [Oryza sativa]"
join(15081. 15228,15380. 15470,15661. >16153)
/gene="F22H5.1"
join(15171. 15228,15380. 15470,15661. 16153)
/gene="F22H5.1"
/codon_start=1
/product="GSH-dependent dehydroascorbate reductase 1,
putative; 14887-15869"
/protein_id="AAG12679.1"
/db_xref="GI:10092266"
/translation="MALDICVKVAVGAPDVLGDCPFQSRVLLTLEEKLPYKTHLINV
SDKPQWFLDISPEGKVPVVKLDGKWVADSDVIVGLLEEKYPEPSLKTTPPEFASVGSKI
FGAFVTFLLSKDANDGSEKALVDELEALHKLTHSGPFVAGEKITAVDLSLAPKLIH
LEVALGHYKNWSVPESLTSVRNYAKALFSRESFENTKAKKEIVVAGWESKVA"
complement(16410. 18554)
/gene="F22H5.2"
complement(join(<16410. 16558,16703. 16830,17050. >18554))
/gene="F22H5.2"
complement(join(16410. 16558,16703. 16830,17050. 18554))
/gene="F22H5.2"
/codon_start=1
/product="unknown protein; 18270-16126"
/protein_id="AAG12682.1"
/db_xref="GI:10092269"
/translation="MPPFRFRIPFFSSNSPSRISSTSPSPPTPPSRPPFRPAGIA
```

Alignment Scores:

```
Pred. No.:      5,46e-72      Length:      68041
Score:          1117.50      Matches:      340
Percent Similarity: 38.10%      Conservative: 73
Best Local Similarity: 31.37%      Mismatches: 82
Query Match:      40.21%      Indels:      595
DB:              8          Gaps:      18
```

US-10-051-909-36 (1-553) x AC025814 (1-68041)

```
QY      52  GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer 71
      |||||:::  ::|||
Db      33823 GAGGAGGCTAGGAATGAT-----TTGCGTCGGCCGTTTATTACACCGGAAGC 33870

QY      72  TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91
      |||||:::  |||||
Db      33871 TGGTACCGGATGGGTTTCGAGACAATCGAGTATG-----ATGGGTCCTCTCAAAGTT 33921

QY      92  LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111
      ::|||:::  ::|||
Db      33922 ATTCGAGACAGCTCTATCTCTGTTCTTGTCATGTGTTTGATCGTTGCTCTTGGCCCTATT 33981

QY      112  GlnPheGlyPheThr----- 116
      |||||:::  |||||
Db      33982 CAATTCGGATTCACTGTAAAGACAAATTCAAAATCTCAAACCTTTTATTATTCCTCC 34041

QY      116  ----- 116
Db      34042 TTTGTAGAAGAAAGGTGTGACTTTTGGTGGAATACTAGTGTGAGCTGGATTGTATC 34101

QY      116  ----- 116
Db      34102 TTGGGTTTGGTTTGATTGTGTCATTTTCAATGCAAGTCTGGTGATTTTGAAAATCTCATATT 34161

QY      116  ----- 116
Db      34162 TGGTGAACCGAGATGATTTTTCAAAGTTTGTCTGACAGATTTAGAGGACTTTGAATTGGGA 34221

QY      116  ----- 116
Db      34222 ATTTCTGATGTTTGGCTTAAACCTTCCATGTTTGACTTTGATCACCATCTTTTAGTACTC 34281

QY      117  -----SerGlyPheSerSerProThrG1 124
      |||||:::  |||||
Db      34282 TGTGTTTGGTCCTTAAACCCTCCATCCTTTTGTTCAGTGTGGTTATTATCATCTCCAACCTCA 34341
```

```
QY      124  nAspAlaMetValArgAspLeuAsnLeuSerIleSerGlu----- 137
      |  |||:::  ::|||
Db      34342 AGCTGCAATCACCAAGGATCTCGGTTTAACAGTATCTGAGGTAATGGAAGTATCATTTGTC 34401

QY      137  ----- 137
Db      34402 TAGAGCTGTCTGGTTTTTATCATGTCTTGTAGCCGCTGTTTGACTTTGATGTACCACTC 34461

QY      138  -----PheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaI1 154
      :  ::|||  |||||
Db      34462 CATGTTTTAGTACTCTGTGTTTGGTTCTCTTTCCAATGTGGTGCTATGGTTGGTCAAT 34521

QY      154  eAlaSerGlyGlnMetAlaGluTyrIleGlyArg-LysGly----- 167
      |||||:::  ::|||
Db      34522 TGCCAGTGGTCAGATTGCAGATACATTTGGACGAAAGGGGTATATACTTTGTTATCTGATT 34581

QY      167  ----- 167
Db      34582 CTTGATCGTTCTGTTTTTAGATCCATTTTATAAAGTCCTCTTTCTTTCGTTCTTAAACATCT 34641

QY      168  -----SerLeuMetIleAlaAlaIleProAsnIleIleGlyTyrL 181
      |||||:::  |||||
Db      34642 TATTGTGTTTTCTTGGACAGTCTCTTATGATTGCTGCAATTCCCAATATCATTTGGATGGC 34701

QY      181  euAlaIleSerPheAlaLys----- 187
      ||  |||||  |||
Db      34702 TTTGCATATCATTTGTC-AAAAGTGAGTTTTTTGTGCTTCTCATACCATAAAATGGCGGATTT 34760

QY      188  -----AspAlaSerPheL 192
      |||||:::  |||||
Db      34761 GATTGTGTTGCTTGACTTTTAACTCTTTTACATCATTTCTGGTTTGTAGGATACCTTCTTTC 34820

QY      192  euTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr----- 209
      |||||:::  |||||
Db      34821 TTTACATGGGAAGACTTTTAGAAGGCTTTGGCGTTGGGATCATCTCTTACACTGTAATAC 34880

QY      209  ----- 209
Db      34881 TGAACTTAACTAGCTATTACTTTATGGAATAACGGTTTCAGTATATTAATGCAACTCGGTT 34940

QY      210  -----ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaL 226
      |||||:::  |||||
Db      34941 CGTGATTTTAGGTACCTGTATATATCGCTGAGATAGCTCCCAAAAATATGAGAGGAGAT 35000

QY      226  euGlySerValAsnGlnLeu----- 232
      |||||:::  |||||
Db      35001 TGGGTTTCGGTTAACCCAGGT-TAGACATGTGTAGTAGCTAATTCCTGTTTGTGCTTTGTG 35059

QY      233  -----SerValT 235
      |||||:::  |||||
Db      35060 TACTGACTCGTAATGCTAAACAAGAAATGTTGTGTAACCTTCTGCTGTAGCTTTCTGTGA 35119

QY      235  hrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlav 255
      |||||:::  |||||
Db      35120 CAATTGGAATAATGTGGCGTATTACTTGGCCCTCTTTGTTCCATGGAGAAATCTTTGCAG 35179

QY      255  al----- 255
      |||||:::  |||||
Db      35180 TTCTGGGTAAATCTCATTCACGGCCTGAGACTATTTTGTGATTATGAATAATTGCTAC 35239

QY      256  -----IleGlyAlaLeuProCysThrMetL 264
      :  ::|||  |||||
Db      35240 GCTCTAATCACTTTCTCTGTTCTCTTCCCTTATCTTTAGGATATTTGCCGTGTACATTAT 35299

QY      264  euIleProGlyLeuPhePheIleProGluSerProArgTyrLeu----- 278
      |||||:::  |||||
Db      35300 TGATACCAGGTCCTTTTTTTCATCCCCGAATCCCTCGCTGGCTGGTGAAGCCTTACAGTAG 35359

QY      278  ----- 278
Db      35360 TTTGCTTATTCTTCTTGAGTTTAGTGCAATTTTGGTTACTCCACTCTTTTTTTTTTTTTT 35419
```



QY 279 -----AlaLysM 281  
Db 35420 TAATAACATCATTAATCATCTCTGATGGGAFACCTTGATTTCGTTAAATGTAGGCAAAA 35479  
QY 281 etAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspI 301  
Db 35480 TGGGTATGACAGATGAATTTGAAACTTCATTACAAGTTCTTCGAGGATTTGAGACTGATA 35539  
QY 301 leThrThrGluValAsnAspIleLys----- 309  
Db 35540 TTACCGTTGAAGTTAATGAAATCAAGGTAACACTACCACCTAAAAATCTTCCTTCCAAGA 35599  
QY 310 -----Arg 310  
Db 35600 GTTACATGAAGACTGGTTTTCAAACGGTAATTTTTCATCTTTCTGCAAAATCTGCAGAGA 35659  
QY 311 AlaValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLys 330  
Db 35660 TCTGTGGCATCATCTACAAAGCGAAATACAGTTTCGGTTTGTAGATCTCAAAACGCAGGAGA 35719  
QY 331 TyrArgThrProLeuLeu----- 336  
Db 35720 TACTATTTCCCACTTATGGTACAAATATTGAATATCTTTCCCTGGAAATATACTTTGTCTA 35779  
QY 337 -----LeuG1 338  
Db 35780 TCATAAGAAACATTATCTTAAAACTTCTTTTGTGTCGTGTGTGTGTAAACAGTTGG 35839  
QY 338 YileGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlase 358  
Db 35840 TATAGGGTTGCTTGTACTTCAACAACCTTGGTGGAAATTAATGGTGTCTTATCTATTCCAG 35899  
QY 358 rSerIlePheLysAlaAla----- 364  
Db 35900 TACAATATTTGAATCTGCAGGTAGGGTTTTTCTAACACCCTGAAGCTTCTTGTTCATATT 35959  
QY 365 -----GlyValThrAsnSerAspLeuA 372  
Db 35960 ATTACTTTAAATTTTTCATTTTCCGTGTTTTCATGTTTAGGAGTTACATCGAGTAACGCAG 36019  
QY 372 laThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuA 392  
Db 36020 CAACATTTGGGGTCGGGGCTATTTCAGGTAAGTAGCGACTGCAATATCCACTTGGTTGGTGG 36079  
QY 392 spArgAlaGlyArgArgIleLeuLeuIle----- 401  
Db 36080 ACAAGCAGGTCGTCGGCTTCTGCTTACTGTGAGTATTAACCTCATTTACTTTTCATATCTTT 36139  
QY 401 ----- 401  
Db 36140 CCGACTTCCCTCTACAGAAAGATTTTAAGACATATCCAAGATTAACTAATGCAAAATGGAAA 36199  
QY 402 -----IleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValp 418  
Db 36200 ATATTTTCCAGATCTCTTCGGTTGGGATGACGATTAGCCTTGTAATGTGTGCAGCTGCTT 36259  
QY 418 hePheLeuLys----- 421  
Db 36260 TCTATCTTAAGGTACTTTTTTAAATCTTGAACTTCCATACACATCATTTCTCTTGCA 36319  
QY 422 -----AspAsnIleSerG 426  
Db 36320 ACAAATCCAAGATTCGACTTTCTGAGTTTACATCTCTTCTTTTAGGAATTT-GTGTCTC 36378  
QY 426 lnAspSerAsnSerTyrTyriIleLeuThrMetIleSerLeuValGlyIleVal----- 443  
Db 36379 CTGATTCAGACATGTACAGTTGGCTGAGCATATTGTTCAGTAGTTGGAGTTGTGGTAAGTT 36438  
QY 443 ----- 443  
Db 36439 AGTTATATAGCGGAACCTTTCTAGAAATGCAATGAATGATTATTTGTGSCCTTAACTCTGTTG 36498  
QY 444 -----SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuM 460

Db 36499 AATCGTTATAGGCAATGGTTGTCTTTTCTCATTTGGGAATGGGACCAATACCGTGGCTCA 36558  
QY 460 etMetSerGluIle----- 464  
Db 36559 TTATGTCTGAGGT-AATGAGTTGGTTAGAAAGTTAGAACTAAGTGTGTGAAGAACCGAAAA 36617  
QY 465 -----L 465  
Db 36618 GCGAAAAAATTTATTGAACCTCTACCTTTGCTTGTCTTGTGTTGTGTTGATAAAACAGATCC 36677  
QY 465 euProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrS 485  
Db 36678 TTCTCTGTGAACATAAAGGGTTTAGCTGGAAGTATTGCAACTCTAGCCAATTGGTTCTTTT 36737  
QY 485 erPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheL 505  
Db 36738 CTTGGTTGATCACCATGACAGCAAAATTTGCTGTTAGCCTGGAGCAGTGGAGGTTCCTTTC 36797  
QY 505 euSerTyrMet----- 508  
Db 36798 TCTCTTTCTC-TTCTTTTCTCTTCTCTCATACCATTGTTGGGACAATAGAGCATTATCTTG 36856  
QY 509 -----ValValSerAlap 513  
Db 36857 ATGGATTTTGGGTTAATATTGGACAGGAACTTTTCACTCTGTATGGATTGGTTTGTGCAT 36916  
QY 513 heThrIleValPheValValLeuTrpValProGluThrLysGly\*\*\*AsnSerArgGlyA 533  
Db 36917 TCACAGTGGTGTTCGTGACTCTATGGGTTCTTGAGACCAAAAGSCAA-AACTCTTGAAGAA 36975  
QY 533 spThrIlePheValSerLeuSerIleGlnArgGlnLeuGlnTrpLeuProGluCysLeuS 553  
Db 36976 CTTCAATCCTTGTTCAGATGAACAAATTTGAACAACCTTCATTTCTTTGTACCCCTCTCT 37035  
QY 553 er 553  
Db 37036 CC 37037

Search completed: April 13, 2005, 22:00:41  
Job time : 5725.66 secs

This page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 12:06:34 ; Search time 638.423 Seconds  
(without alignments)  
5127.663 Million cell updates/sec

Title: US-10-051-909-36  
Perfect score: 2779  
Sequence: 1 PSSSSSFRPAGKKKKKQNG.....TIFVSLSIQRQLQWLPECLS 553

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074037\_14040/app\_query.fasta\_1.1678  
-DB=N\_Geneseq\_16Dec04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909 @CGN 1 1 891 @runat\_13042005\_074037\_14040 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :				N_Geneseq_16Dec04:*	
1:	geneseqn1980s:	1	1980	1980	1980
2:	geneseqn1990s:	2	1990	1990	1990
3:	geneseqn2000s:	3	2000	2000	2000
4:	geneseqn2001as:	4	2001	2001	2001
5:	geneseqn2001bs:	5	2001	2001	2001
6:	geneseqn2002as:	6	2002	2002	2002
7:	geneseqn2002bs:	7	2002	2002	2002
8:	geneseqn2003as:	8	2003	2003	2003
9:	geneseqn2003bs:	9	2003	2003	2003
10:	geneseqn2003cs:	10	2003	2003	2003
11:	geneseqn2003ds:	11	2003	2003	2003
12:	geneseqn2004as:	12	2004	2004	2004
13:	geneseqn2004bs:	13	2004	2004	2004

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	DB ID	Description
1	2777	99.9	1953	12	ADG47939 Corn Beta
2	1710	61.5	1377	8	ADA69666 Rice gene
3	1577	56.7	1428	8	ADA69667 Rice gene
4	1182.5	42.6	1077	13	ADR59784 Cotton cD
5	951.5	34.2	1741	3	AAC36319 Arabidops

6	946.5	34.1	990	13	ADR59785	Adr59785 Cotton cD
7	943.5	34.0	1683	3	AAC39099	Aac39099 Arabidops
8	933	33.6	1395	3	AAC45857	Aac45857 Arabidops
9	923	33.2	1685	3	AAC45298	Aac45298 Arabidops
10	923	33.2	1743	3	AAC45290	Aac45290 Arabidops
11	920.5	33.1	1473	6	ABZ14526	Abz14526 Arabidops
12	916	33.0	1755	3	AAC40459	Aac40459 Arabidops
13	899	32.3	1359	6	ABZ13822	Abz13822 Arabidops
14	893	32.1	1768	12	ADM47939	Adm47939 Polynucle
15	877.5	31.6	1407	6	ABZ13486	Abz13486 Arabidops
16	877.5	31.6	1407	8	ADA67841	Ada67841 Arabidops
17	877.5	31.6	1642	3	AAC46009	Aac46009 Arabidops
18	868.5	31.3	1425	3	AAC50050	Aac50050 Arabidops
19	853.5	30.7	1425	3	AAC50049	Aac50049 Arabidops
20	837.5	30.1	1209	8	ADA70113	Ada70113 Rice gene
21	824	29.7	1452	3	AAC42659	Aac42659 Arabidops
22	805.5	29.0	1695	3	AAC49517	Aac49517 Arabidops
23	790.5	28.4	1436	3	AAC45400	Aac45400 Arabidops
24	786.5	28.3	1440	3	AAC33666	Aac33666 Arabidops
25	777	28.0	1374	6	ABZ13215	Abz13215 Arabidops
26	758.5	27.3	1378	3	AAC35123	Aac35123 Arabidops
27	754.5	27.2	1344	8	ADA69777	Ada69777 Rice gene
28	688.5	24.8	1724	4	ABL13389	Ab113389 Drosophil
29	682.5	24.6	1239	3	AAC34742	Aac34742 Arabidops
30	674.5	24.3	2072	4	AAF55867	Aaf55867 Murine GL
31	674	24.3	2217	4	AAF55865	Aaf55865 Human GLU
32	672	24.2	898	10	ADD17645	Add17645 DNA (SeqI
33	672	24.2	2087	4	AAF55866	Aaf55866 Rat GLUTX
34	669.5	24.1	2080	4	AAD09552	Aad09552 Human tra
35	663	23.9	1473	6	ABZ13864	Abz13864 Arabidops
36	663	23.9	1473	6	ADG88039	Adg88039 A. thalia
37	658.5	23.7	1445	8	ABZ24794	Abz24794 Human sol
38	658.5	23.7	1856	8	ABZ24792	Abz24792 Human sol
39	657.5	23.7	600	13	ACN59223	Acn59223 Cotton gy
40	652.5	23.5	1873	8	ABZ24793	Abz24793 Human sol
41	652.5	23.5	2514	4	ABL17067	Ab117067 Drosophil
42	640	23.0	1412	4	ABL13883	Ab113883 Drosophil
43	623.5	22.4	2487	8	ABZ24790	Abz24790 Human sol
44	623.5	22.4	2487	12	ADQ86239	Adq86239 Human tum
45	623.5	22.4	2487	13	ADR25824	Adr25824 Breast ca

ALIGNMENTS

RESULT 1	
ADG47939	
ID	ADG47939 standard; cDNA; 1953 BP.
XX	
AC	ADG47939;
XX	
DT	11-MAR-2004 (first entry)
XX	
DE	Corn Beta-vulgaris-like sugar transport protein cDNA #4.
XX	
KW	Arabidopsis-like sugar transport protein;
KW	Beta-vulgaris-like sugar transport protein; transgenic; physical mapping;
KW	corn; plant; gene; ss.
XX	
OS	Zea mays.
XX	
FH	Key
CDS	Location/Qualifiers
FT	1..1662
FT	/*tag= a
FT	/product= "Corn Beta-vulgaris-like sugar transport
FT	protein"
FT	/note= "Xaa can be any amino acid, CDS does not include
FT	no start codon"
FT	/transl_except= (pos:1582..1584, aa:Xaa)
FT	/partial
FT	
XX	
PN	US2002199217-A1.
XX	
PD	26-DEC-2002.



XX	17-JAN-2002; 2002US-00051909.
PF	
XX	
PR	24-APR-1998; 98US-0083044P.
PR	14-APR-1999; 99US-00291922.
XX	
PA	(HELE/) HELENTJARIS T G.
XX	
PI	Helentjaris TG;
XX	
DR	WPI; 2004-040967/04.
DR	P-PSDB; ADG47940.
XX	
PT	New isolated polynucleotide encoding a polypeptide having sugar transport
PT	protein activity, for producing a transformed plant and for use as probes
PT	in physical mapping.
XX	
PS	Claim 6; SEQ ID NO 35; 71pp; English.
XX	
CC	The invention relates to Arabidopsis-like or Beta-vulgaris-like sugar
CC	transport proteins and their corresponding nucleic acid sequences. The
CC	sequences of the invention are useful to transform a cell. These are also
CC	useful to produce a transgenic plant. Probes derived from sequences
CC	encoding sugar transport protein may be used for physical mapping. The
CC	present sequence is corn Beta-vulgaris-like sugar transport protein cDNA.
XX	
SQ	Sequence 1953 BP; 466 A; 478 C; 474 G; 534 T; 0 U; 1 Other;

Qy	181	LeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGly	200
Db	541	CTTGGGATCTCCTTTGCAAAAGATGCCTCATTTCTATATATGGACGATTGCTTGAAGGG	600
Qy	201	PheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGln	220
Db	601	TTTGGTGTGGCATCATATCTTACACGGTACCGGTATACATAGCAGAGATATCTCCTCAG	660
Qy	221	AsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeu	240
Db	661	AACATGAGGGGAGCTCTTGGTCTGTGAACCAAGTTGCTGTGACCTTTGGCATATTCTTG	720
Qy	241	AlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuPro	260
Db	721	GCCTATTGCTCGGCATGTTATTCTTTGGAGACTTCTTGCTGTGATTGGAGCCCTTGCCC	780
Qy	261	CysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLys	280
Db	781	TGCACAAATGTTGATTCCTGGACTATTCTTCAATCCAGAATCTCCAGATGGCTGGCAAG	840
Qy	281	MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAsp	300
Db	841	ATGAATTTGACCGGAAGATTGTGAGACGTCCCTACAAAGTGTGAGGGGGTTTGAGACTGAC	900
Qy	301	IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThr	320
Db	901	ATCACAAACAGAAGTGAATGATATAAAGAGGGCAGTGGCATCATCAAGTAAGAGGACCACA	960
Qy	321	IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGly	340
Db	961	ATCAGTTTTCAGAATTAACCAAAAGAAATACCGCACGCCACTACTTCTAGGGATTGGC	1020
Qy	341	LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIle	360
Db	1021	CTACTTGTACTGCAAAATCTTAGTGAATCAACGGTGTACTGTTTTATGCAAGTAGCATC	1080
Qy	361	PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln	380
Db	1081	TTCAAAGCTGCAGGGGTACAAACAGCGACTTGGCCACCCTGTTTCACTTGGTGTATTTCAG	1140
Qy	381	ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu	400
Db	1141	GTCCTTGCTACTGGAGTTACAACATGGCTGTAGACCGAGCTGGACACGCATCCTTCTC	1200
Qy	401	IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeu	420
Db	1201	ATTATTCTACCTCTGGCATGACTCTATGCGCTTCTTGCCGTTTCTGTTGTAATTTTCTC	1260
Qy	421	LysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuVal	440
Db	1261	AAGGATAACATTTACAGGATTCTAACTCATACTACATCTTAACAAATGATCTCCCTTGT	1320
Qy	441	GlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMet	460
Db	1321	GGTATTGTGTCCTTTGTCTATTACCTTCTCGTTTGGTATGGTGCCATTCACATGGCTCATG	1380
Qy	461	MetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAla	480
Db	1381	ATGTCCTGAGATCCTCCCGGTTAGCATCAAGAGCCTTGGCGGAAGCATCGCAACACTGGCC	1440
Qy	481	AsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVal	500
Db	1441	AACCTGGCTGACATCCTTCGCCATAACAAATGACGACGAACCTTGATGCTCACGTGGAGTGT	1500
Qy	501	GlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeu	520
Db	1501	GGAGGCATTTTCTCTCGTACATGGTTGTGAGCGCCTTCACCATCGTTTTTGTGTCTCT	1560
Qy	521	TrpValProGluThrLysGly***AsnSerArgGlyAspThrIlePheValSerLeuSer	540
Db	1561	TGGTGCCGGGACGAAGGGGAGNACTCTAGAGGAGATACAAATTTTCGTTTCGCTGAGC	1620

QY 541 IleGlnArgGlnLeuGlnTrpLeuProGluCysLeuSer 553  
Db 1621 ATTCAGCGTCAGTGCAATGGTTGCCCGAGTGTATCT 1659

RESULT 2  
ADA69666  
ID ADA69666 standard; DNA; 1377 BP.  
XX  
AC ADA69666;  
XX

DT 20-NOV-2003 (first entry)  
DE Rice gene, SEQ ID 2989.  
XX  
KW Plant; bacterial infection; fungal infection; viral infection; rice;  
KW gene; ds.  
XX  
OS Oryza sativa.  
XX  
PN WO2003000898-A1.  
XX  
PD 03-JAN-2003.  
XX

XX 22-JUN-2001; 2001WO-IB001105.  
XX 22-JUN-2001; 2001WO-IB001105.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.

XX Claim 6; SEQ ID NO 2989; 899pp; English.  
PS The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

XX Sequence 1377 BP; 290 A; 338 C; 360 G; 388 T; 0 U; 1 Other;

Alignment Scores:  
Pred. No.: 1.95e-145 Length: 1377  
Score: 1710.00 Matches: 358  
Percent Similarity: 78.36% Conservative: 44  
Best Local Similarity: 69.79% Mismatches: 39  
Query Match: 61.53% Indels: 73  
DB: 8 Gaps: 9

US-10-051-909-36 (1-553) x ADA69666 (1-1377)

QY 44 AsnArg---GlyGlyAlaGlyAlaGlyGluSerGlySerAspHisAasp-----Gly 60  
Db 4 AACAGGGCGGCGGCGTCGGCGTGAGAGCGGAAGCGACTACGAGAGCGGCGGC 63  
QY 61 ValLeuArgArgProLeuLeu---AsnThrGlySerTrpTyrArgMetSerSerArgGln 79  
Db 64 GGCATGCGGAAGCGCGTCTGATGCACACGCGGAGCTGGTACAGGATGGGGTCGCGGCAG 123  
QY 80 SerSerPhe---AlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98

Db 124 GGGAGCCTCACCGGCGGGGACCTCGTCATGGCCATCCTACGGAGTCCCACGTCTCC 183  
QY 99 AlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGly 118  
Db 184 GCCTTCCTCTGCACGCTCATCGTCGCGCTCGGCCCCATCCAGTTCGAGTTCACGGGGGC 243  
QY 119 PheSerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPhe 138  
Db 244 TTCTCCTCCCGACGACGAGCGCCATCATCCGAGACCTCGACCTCACCTCTCCGAGTTC 303  
QY 139 SerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGln 158  
Db 304 TCGGTGTTGGATCGCTGTCCAACGTCGGGCCCATGGTTGGGGCGATTGCCAGTGGTCAA 363  
QY 159 MetAlaGluTyrIleGlyArgLysGly----- 167  
Db 364 ATGCCCGAGTACATTGGGGCGAAAGGKTGAGAAATTTTTCCTTTTGGGTACACATC 423  
QY 168 SerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys 187  
Db 424 TCATTGATGATTGCTGCAATTCCAAACATCATTTGGTTGGCTTGCCATCTCCTTTGCAAAA 483  
QY 188 AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSer 207  
Db 484 GACTCATCGTTTCTTATATGGGACGATTGCTCGAGGGGTTTGGTGTGGTGTCACTCTCT 543  
QY 208 TyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGly 227  
Db 544 TATACGGTGCCAGTTTACATAGCAGAAATATACCTCAAAACATGAGAGGTGCTCTTGGC 603  
QY 228 SerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe 247  
Db 604 TCAGTGAATCAGTTATCTGTAAACCGTTGGTATATTGTTGGCATATTTGCTCGGCATGTTT 663  
QY 248 IleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267  
Db 664 GTTCCTTGAGGCTTCTTGCTGTATAGGAATCTTGCCCTTGCACTGTGTGTATACCTGGC 723  
QY 268 LeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCys 287  
Db 724 CTATTCTTCATTCCAGAAATCCCAAGATGGTTGGCAAAGATGAACATGATGGATGATTTT 783  
QY 288 GluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAsp 307  
Db 784 GAGACTTCTTTACAAGTCTGAGGGGATTTGAGACTGACATCAGCGCGGAAGTGAATGAT 843  
QY 308 IleLysArgAlaValAlaSerSerSerLysArgThrThrIleSer---PheGlnGluLeu 326  
Db 844 ATAAAG-----GTAGCATCTTCAAAGCAGCAGTCATATTTTAGCCGATTTGGCTTCGTG 897  
QY 327 AsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsn 346  
Db 898 ACACGCAAA----- 906  
QY 347 LeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyVal 366  
Db 907 -----ATTTCTTCAGGTCTC 921  
QY 367 ThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyVal 386  
Db 922 ACAAAACAGTGACTTGGCTACATGTGCACCTGGTGGTGTATCCAG----- 963  
QY 387 ThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGly 406  
Db 964 -----ATCTCTTCTGCTGGG 978  
QY 407 MetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGln 426  
Db 979 ATGACTCTAAGCCTCCTTGCAGTTGCTGTGTATTTTTCCTCAAGGATAGCATTTCACAA 1038  
QY 427 AspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheVal 446

Db 1039 GATTCTCACATGTACTACACCTTTAAGTATGATCTCCTTGGTTGCTCTTGTGGCTTTTGTA 1098  
QY 447 IleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuPro 466  
Db 1099 ATCGCCTTCTCCTTCGGTATGGGTGCCATTCCATGGATCATAATGTCAGAGATCCTCCCG 1158  
QY 467 ValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPhe 486  
Db 1159 GTTAGTATCAAGAGTCTCGCAGGAAGCTTTGCGACGCTCGCCAACTGGCTTACATCCTTT 1218  
QY 487 AlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSer 506  
Db 1219 GGAATAACAATGACAGCAAACTTGATGCTTAGCTGGAGTGCTGGAGGGACCTTTGTGTCC 1278  
QY 507 TyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThrLys 526  
Db 1279 TACATGGTTCGTGAGTGCTTTTCAACCCTCGTGTTCGTCATCCTTTGGGTGCCAGACAAAA 1338  
QY 527 Gly\*\*\*AsnSerArgGlyAspThrIlePheValSerLeu 539  
Db 1339 GGAAG-AACTCTCGAAGAGATACAAATGGTCCTTCCGCTG 1376

RESULT 3

ADA69667

ID ADA69667 standard; DNA; 1428 BP.

XX ADA69667;

DT 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 2990.

DE Plant; bacterial infection; fungal infection; viral infection; rice;  
KW gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

PF 22-JUN-2001; 2001WO-IB001105.

PR 22-JUN-2001; 2001WO-IB001105.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.

XX Claim 6; SEQ ID NO 2990; 899pp; English.

CC The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

XX Sequence 1428 BP; 352 A; 315 C; 354 G; 407 T; 0 U; 0 Other;

SQ Alignment Scores:

Pred. No.: 2.57e-133 Length: 1428  
Score: 1577.00 Matches: 318  
Percent Similarity: 85.85% Conservative: 34  
Best Local Similarity: 77.56% Mismatches: 32  
Query Match: 56.75% Indels: 27  
DB: 8 Gaps: 2  
US-10-051-909-36 (1-553) x ADA69667 (1-1428)  
QY 130 AspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGly 149  
Db 277 GATCTGTACTTACAACAGCATATGTTCTCGGCTTTCGGTTCGCTGTCCAACGTCGGCGCC 336  
QY 150 MetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeu 169  
Db 337 ATGGTCGGAGCGATCGCCAGTGGGCAGATGGCGGAGTACATTGGCCGGAAGGTCGTTG 396  
QY 170 MetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAla 189  
Db 397 ATAATTGCGCGGTTCTTAACATCATTTGGTTGGCTTGCATCTCCTTTGCAAAAGACGCG 456  
QY 190 SerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209  
Db 457 TCATTTTATACATGGGACGCTTGCTTGAAGTTTGGTGTGGTGTATATCATATACG 516  
QY 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229  
Db 517 GTGCCAGTATACATAGCAGAGATATCTCATCAGAACACAGAGGAGCACTTGGCTCCGTG 576  
QY 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIlePro 249  
Db 577 AACCAGTTGTCGTTACCATGGTATCTTGTGGCTATTTGTAGGCATGTTTGTTCCT 636  
QY 250 TrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhe 269  
Db 637 TGGAGGCTGCTTGCACTGATAGGAAGCATCCCATGTACATTGTTTATACCTGGTCTATTC 696  
QY 270 PheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThr 289  
Db 697 TTCATTCCGGAATCCCAAGATGGCTGGCAAAAATGAAAATGATGGATGATTTTGAGGCT 756  
QY 290 SerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLys 309  
Db 757 TCTCTACAAGTTTGTAGGGGTTTGAGACTGATATATTACTGCAGAA----- 801  
QY 310 ArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLys 329  
Db 802 AGAGCAGTAGCATCGGCAAAACAAAAGAACCCACAGTCCGTTTAAAGAGTTGAACCAAAAG 861  
QY 330 LysTyrArgThrProLeuLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGly 349  
Db 862 AAATACCGCACTCCCTTACTGATAGGAACCTGGCCTTCTTGTACTTCAGAAATCTAAGTGA 921  
QY 350 IleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSer 369  
Db 922 ATAAATGGTATTCTGTTTATGCAAGTAGAATCTTCAGAGATGCAGGGTTTACAAACAGT 981  
QY 370 AspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrp 389  
Db 982 GACTTGGCCACATGTGCACCTAGGAGCAATTCAG----- 1014  
QY 390 LeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeu 409  
Db 1015 -----ATCTCTACTGCTGGGATGACTCTA 1038  
QY 410 CysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsn 429  
Db 1039 AGCCTTCTTGCAGTATCTGTTGATTTTTCCTGGAGGGTAATATTTTCACATGATTCAT 1098  
QY 430 SerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPhe 449  
Db 1099 TCGTTCTACATCTTAAGTATGATCTCCTTGGTTGCTCTTGTGGCTTATATCATCACCTTT 1158





Db	538	GTGTATATAGCTGAGATAGCACCCGAAAACTTGAGAGGCAGCTTGGGTTCTGTGAATCAG	597	PR	04-MAY-1999;	99US-0132484P.
QY	232	LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTirArg	251	PR	05-MAY-1999;	99US-0132485P.
Db	598	TTGTCIGTTACAACTGGAACAATGGTAGCCCTATTGCTGGGACTTTTGTGGGTGGAGG	657	PR	06-MAY-1999;	99US-0132486P.
QY	252	LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle	271	PR	07-MAY-1999;	99US-0132487P.
Db	658	ATACTTGAATTTTAGGAGGTGTGCTTGACAAATTTGTATCCAGGCTATTTTTCATT	717	PR	11-MAY-1999;	99US-0134256P.
QY	272	ProGluSerProArgTirPLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu	291	PR	14-MAY-1999;	99US-0134218P.
Db	718	CCAGAACTCTCCTAGATGGCTGGCAAAAAATGGGAATGACCGAAGATTTCGAAGCTTCTCTT	777	PR	14-MAY-1999;	99US-0134221P.
QY	292	GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla	311	PR	14-MAY-1999;	99US-0134370P.
Db	778	CAAGTTCCTTAGGGGTTTCGATACCGATATTTCTATCGAAGTCAATGAAATCAAGAGGTCT	837	PR	18-MAY-1999;	99US-0134768P.
QY	312	ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLys-LysTy	331	PR	19-MAY-1999;	99US-0134941P.
Db	838	GTAGCATCAACAAATAGAGAACAACACGATCCCGTTCCGAGAACTCAAAAAAAGGAGATA	897	PR	20-MAY-1999;	99US-0135124P.
QY	331	rArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAs	351	PR	21-MAY-1999;	99US-0135353P.
Db	898	TTGGTTCCATTGATGGTTGGAATTGGGCTTCTATGTTGCAACAACCTAGTGGCAATTAA	957	PR	24-MAY-1999;	99US-0135629P.
QY	351	nGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLe	371	PR	25-MAY-1999;	99US-0136021P.
Db	958	TGGTGTATATTCTATTCGCTACCATATTCGAAACTGCTGGAATTAATCAGGCAATAT	1017	PR	27-MAY-1999;	99US-0136392P.
QY	371	uAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTirPLeu	390	PR	28-MAY-1999;	99US-0136782P.
Db	1018	AGCTACCTTTGGAGTTGGCTTCATTCAAGTTCATGCTACTGCTTTAAACCACATGGTTG	1075	PR	01-JUN-1999;	99US-0137222P.
RESULT 5						99US-0137502P.
ID	AAC36319 standard; DNA; 1741 BP.					99US-0137724P.
XX	AAC36319;					99US-0138094P.
AC	AAC36319;					99US-0138540P.
XX	17-OCT-2000 (first entry)					99US-0138847P.
DT	Arabidopsis thaliana DNA fragment SEQ ID NO: 13356.					99US-0139119P.
DE	Hybridisation assay; genetic mapping; gene expression control;					99US-0139452P.
XX	protein identification; signal transduction pathway; metabolic pathway;					99US-0139453P.
KW	promoter; termination sequence; ss.					99US-0139454P.
XX	Arabidopsis thaliana.					99US-0139455P.
OS	EP1033405-A2.					99US-0139456P.
PN	06-SEP-2000.					99US-0139457P.
XX	25-FEB-2000; 2000EP-00301439.					99US-0139458P.
PF	25-FEB-1999;					99US-0139459P.
XX	05-MAR-1999;					99US-0139460P.
PR	09-MAR-1999;					99US-0139461P.
PR	23-MAR-1999;					99US-0139462P.
PR	25-MAR-1999;					99US-0139463P.
PR	29-MAR-1999;					99US-0139750P.
PR	01-APR-1999;					99US-0139817P.
PR	06-APR-1999;					99US-0139899P.
PR	08-APR-1999;					99US-0140353P.
PR	16-APR-1999;					99US-0140354P.
PR	21-APR-1999;					99US-0140695P.
PR	23-APR-1999;					99US-0140823P.
PR	28-APR-1999;					99US-0140991P.
PR	30-APR-1999;					99US-0141287P.
PR	25-FEB-1999;					99US-0141842P.
PR	05-MAR-1999;					99US-0142154P.
PR	09-MAR-1999;					99US-0142390P.
PR	23-MAR-1999;					99US-0142803P.
PR	25-MAR-1999;					99US-0142920P.
PR	29-MAR-1999;					99US-0142977P.
PR	01-APR-1999;					99US-0143542P.
PR	06-APR-1999;					99US-0143624P.
PR	08-APR-1999;					99US-0144005P.
PR	16-APR-1999;					99US-0144085P.
PR	21-APR-1999;					99US-0144325P.
PR	23-APR-1999;					99US-0144331P.
PR	28-APR-1999;					99US-0144332P.
PR	30-APR-1999;					99US-0144333P.
PR	25-FEB-1999;					99US-0144334P.
PR	05-MAR-1999;					99US-0144335P.
PR	09-MAR-1999;					99US-0144352P.
PR	23-MAR-1999;					99US-0144632P.
PR	25-MAR-1999;					99US-0144884P.
PR	29-MAR-1999;					99US-0144814P.
PR	01-APR-1999;					99US-0145086P.
PR	06-APR-1999;					99US-0145088P.

PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.

PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Alignment Scores: 2.68e-76 Length: 1741  
Pred. No.: 951.50 Matches: 181  
Score: 66.12% Conservative: 102  
Percent Similarity: 42.29% Mismatches: 140  
Best Local Similarity: 34.24% Indels: 5  
Query Match: 3 Gaps: 2  
DB: 3  
  
US-10-051-909-36 (1-553) x AAC36319 (1-1741)  
  
Qy 100 PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe 119  
Db 287 TATCTCAGCACATTTCGTTGCGGTCTGTGGTCCTTTTGGGTTTCGGTCTTGTGCGGGATAC 346  
  
Qy 120 SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer 139  
Db 347 TCttCACCTGCTCAGGCTGCAATTAGGAATGACCTTTTCATTGACTATAGTGTGAGTTTTCa 406  
  
Qy 140 AlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet 159  
Db 407 CTCttTGGTtCTTTACTAACTTTTGGCGCAATGATCGGTGCTATACAAGCGGGCCTATA 466  
  
Qy 160 AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179  
Db 467 GCTGATTAGTTGGAAGAAAGGGGGCGATGAGAGTTTCCTCTGCATTTTGTGTAGTCGGG 526  
  
Qy 180 TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu 199  
Db 527 TGGCTAGCAATCATCTTTGCCAAGGGAGTAGTGGCTCTGGACCTCGGAAGACTGGCAACG 586  
  
Qy 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219  
Db 587 GGATATGGAATGGGAGCATTTTCCTATGTGGTGCCAATCTTTATAGCAGAAATTGCACT 646  
  
Qy 220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe 239  
Db 647 AAAACTTTCAGAGGGGCTCTAACCAACACTGAACCAGATTCTGATCTGCACTGGAGTGTG 706  
  
Qy 240 LeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeu 259  
Db 707 GTTTCCTTCATCATAGGCACACTAGTGACGTGGAGAGTCTTGGCATTAATAGGAATCATC 766  
  
Qy 260 ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAla 279  
Db 767 CCATGCGCTGCCCTCCTTCCTTGGGCTCTTTTTTATCCCTGAGTCTCCAAGATGGCTGGCA 826  
  
Qy 280 LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr 299  
Db 827 AAAGTGGGGCGTGATACGGAGTTTGAAGCTGCACGTAGGAAGCTCCGTGGGAAGAGGCT 886  
  
Qy 300 AspIleThrThrGluValAsnAspIleLysArgAlaValaLaserSerSerLysArgThr 319  
Db 887 GATATTTTCGGAGGAGGCAGCAGAGATCCAGGATTATATCGAAACTCTTGGAAAGGCTACCG 946  
  
Qy 320 ThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIle 339



Db 947 AAAGCCAGATGCTGGATTTGTTTCAGAGGAGATACATACGCTCTGTCTTATAGCTTTC 1006  
Qy 340 GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer 359  
Db 1007 GGGTTGATGGTGTTCAGCAAGTTTGGAGGAATCAACGGAATATGTTTCTACACAAGCTCG 1066  
Qy 360 IlePheIysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle 379  
Db 1067 ATATTTGAGCAAGCAGGTTTTCCTCC--ACAAGACTGGGATGATATATATGCTGTTCTT 1123  
Qy 380 GlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeu 399  
Db 1124 CAGGTGTAATCACTCGCTTAATGCACCGATAGTTGACAGAGCCGGAAGAAACCATTTG 1183  
Qy 400 LeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePhe 419  
Db 1184 CTACTGGTTCTGCAACAGGGTAGTGATAGGCTGTTTGATAGCAGCGGTTTCTTTCTAT 1243  
Qy 420 LeuIysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeu 439  
Db 1244 CTCAGGTTTCAC-----GACATGGCGCACGAAGCAGTCCAGTCCCTGGCTGTT 1291  
Qy 440 ValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeu 459  
Db 1292 GTTGGTATAATGGTGATACATAGGATCGTTTTTCAGCAGGAATGGGAGCAATGCCGTGGGTG 1351  
Qy 460 MetMetSerGluIleLeuProValSerIleIysSerLeuGlyGlySerIleAlaThrLeu 479  
Db 1352 GTCATGTCTGAGATATTCCCATAAACATAAAAGGAGTAGCAGCAGGCATGGCGACGCTG 1411  
Qy 480 AlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSer 499  
Db 1412 GTGAACCTGGTTTGGAGCGTGGCTGTTTCTTACACTTTCAACTTCCTCATGTCTCGGAGC 1471  
Qy 500 ValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValVal 519  
Db 1472 TCTTACGGAACCTTCCTCATTTACGCTGCCATCAACGGCGTGGCCATTGTTCTTCGTCATT 1531  
Qy 520 LeuTrpValProGluThrLysGly 527  
Db 1532 GCCATCGTGCCTGAGACAAAAGGG 1555

RESULT 6

ADR59785  
ID ADR59785 standard; cDNA; 990 BP.  
XX  
AC ADR59785;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Cotton cDNA sequence, SEQ ID 566.  
XX  
KW Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;  
KW drought tolerance; plant disease resistance; galactomannan; lignin;  
KW plant growth regulator; heat tolerance; herbicide tolerance;  
KW homologous recombination; extreme osmotic condition tolerance;  
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;  
KW stress resistance.  
XX  
OS Gossypium hirsutum.  
XX  
PN US2004181830-A1.  
XX  
PD 16-SEP-2004.  
XX  
PF 29-JAN-2004; 2004US-00767795.  
XX  
XX 07-MAY-2001; 2001US-00849529.  
PR 12-DEC-2001; 2001US-00021323.  
XX  
XX (KOVA/) KOVALIC D K.  
PA (ZHOU/) ZHOU Y.

PA (CAOY/) CAO Y.  
XX Kovalic DK, Zhou Y, Cao Y;  
PI WPI; 2004-667718/65.  
XX  
PT New recombinant nucleic acid molecules and polypeptides from Gossypium  
PT hirsutum, useful for producing plants with improved biological  
PT characteristics (e.g. improved plant cold or drought tolerance).  
XX  
PS Claim 1; SEQ ID NO 566; 14pp; English.  
XX  
CC The invention relates to a recombinant polynucleotide comprising any of  
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.  
CC Also a recombinant polypeptide comprising any of the 58798 amino acid  
CC sequences mentioned in the specification and producing a plant having an  
CC improved property. Producing a plant having an improved property  
CC comprises transforming a plant with a recombinant construct comprising a  
CC promoter region functional in a plant cell operably joined to a  
CC polynucleotide comprising a coding sequence for a polypeptide associated  
CC with the property, and growing the transformed plant. The polypeptide is  
CC useful for improving plant cold tolerance, manipulating growth rate in  
CC plant cells by modification of the cell cycle pathway, improving plant  
CC drought tolerance, providing increased resistance to plant disease,  
CC producing galactomannan (or lignin or plant growth regulators), improving  
CC plant heat tolerance, improving plant tolerance to herbicides, increasing  
CC the rate of homologous recombination in plants, improving plant tolerance  
CC to extreme osmotic conditions or to pathogens or pests, improving yield  
CC by modification of photosynthesis, modifying seed oil or protein yield  
CC and/or content, improving yield by modification of carbohydrate, nitrogen  
CC or phosphorus use and/or uptake, or improving yield by providing improved  
CC plant growth and development under at least one stress condition. The  
CC polynucleotide and polypeptide may also be used in recombinant DNA  
CC constructs, in physical arrays of molecules, as plant breeding markers,  
CC or in computer-based storage and analysis systems. The present sequence  
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20040181830. However only 6585  
CC polynucleotide sequences were available, the remaining 52213  
CC polynucleotides and all 58798 protein sequences were not present.  
XX  
SQ Sequence 990 BP; 239 A; 192 C; 232 G; 327 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.33e-76 Length: 990  
Score: 946.50 Matches: 196  
Percent Similarity: 80.55% Conservative: 40  
Best Local Similarity: 66.89% Mismatches: 53  
Query Match: 34.06% Indels: 5  
DB: 13 Gaps: 1

US-10-051-909-36 (1-553) x ADR59785 (1-990)

Qy 53 GluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySerTrp 72  
Db 124 GATGATGCTGAAGATGGAAGAGGGGATCTTAAGGAAGCCATTCTTACATACTGGAAGCTGG 183  
Qy 73 TyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValLeu 92  
Db 184 TATCGTATGGGTTCTAGGATGGGGTCTAGTATGTTAGGATCCTCTCAA-----GCTCTT 237  
Qy 93 ArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGln 112  
Db 238 CGTGATAAAATCCATCTCGGTTGTAGCTTGTTTTGATTGTTGTTGGTCCCATCCAA 297  
Qy 113 PheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsn 132  
Db 298 TTTGGTTTCACCTTCGTTACTCTTCCGACACAAATCCGCAATCATCGAGGATCTTGA 357  
Qy 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152  
Db 358 CTAACAGTCTCAGAGTTTCTCTATTGTTGTTCTTTATCAAAATGTGGTGCCATGTTTGA 417

Qy	153	AlaileAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAla	172	PR	19-APR-1999;	99US-0130077P.
				PR	21-APR-1999;	99US-0130449P.
				PR	23-APR-1999;	99US-0130510P.
Db	418	GCAATAGCCAGTGGTCAGATAGCCGAATGATGGACCGAAAGGTCCTTAATGACTGCT	477	PR	23-APR-1999;	99US-0130891P.
				PR	28-APR-1999;	99US-0131449P.
Qy	173	Alaile-ProAsnIleIleGlyTrpLeuAlaIleSerPheAlaIleSerPheLe	192	PR	30-APR-1999;	99US-0132048P.
				PR	30-APR-1999;	99US-0132407P.
Db	478	GCAATTCCCTAAATATAATTGGATGGCTTGCTATATCTTTGCAAGAGATTCTCTTTTCT	537	PR	04-MAY-1999;	99US-0132484P.
				PR	05-MAY-1999;	99US-0132485P.
Qy	192	uTyrMetGlyArgLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProva	212	PR	06-MAY-1999;	99US-0132486P.
				PR	06-MAY-1999;	99US-0132487P.
Db	538	TTACATGGGAAGGCTGTTGCAAGGTTCTCGTGCGCATAAATCTTTATACGGTGCCCTGT	597	PR	07-MAY-1999;	99US-0132863P.
				PR	11-MAY-1999;	99US-0134256P.
Qy	212	lTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLe	232	PR	14-MAY-1999;	99US-0134218P.
				PR	14-MAY-1999;	99US-0134219P.
Db	598	ATATATAGCTGAGATAGCACCTCAGAAATTGAGGGGCAGCTTGGGTTCAGCGAATCAGCT	657	PR	14-MAY-1999;	99US-0134221P.
				PR	14-MAY-1999;	99US-0134370P.
Qy	232	uSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLe	252	PR	18-MAY-1999;	99US-0134768P.
				PR	19-MAY-1999;	99US-0134941P.
Db	658	ATCTGTCACTCTCGGAATAATGCTGGCCTATCTACTTCGACTTTTGTGAGTGCCGGAT	717	PR	20-MAY-1999;	99US-0135124P.
				PR	21-MAY-1999;	99US-0135353P.
Qy	252	uLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIlePr	272	PR	24-MAY-1999;	99US-0135629P.
		:		PR	25-MAY-1999;	99US-0136021P.
Db	718	ACTCGCGCTTTAGGAATACTCTCTGTGACAATTTTGATACCTGCTCTATTTTTCATTCC	777	PR	27-MAY-1999;	99US-0136392P.
				PR	28-MAY-1999;	99US-0136782P.
Qy	272	oGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGl	292	PR	01-JUN-1999;	99US-0137222P.
				PR	03-JUN-1999;	99US-0137528P.
Db	778	AGAATCTCTCGATGGCTGGCAAAACTGGGTATGACAGAAGATTTCGAAACTTCTTTGCA	837	PR	04-JUN-1999;	99US-0137502P.
				PR	07-JUN-1999;	99US-0137724P.
Qy	292	nValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlava	312	PR	08-JUN-1999;	99US-0138094P.
				PR	10-JUN-1999;	99US-0138540P.
Db	838	AGTTCAGGTGCTCCGATGCTGATATTCTATTGAAGTGAATGAAATCAACAGATCTGT	897	PR	10-JUN-1999;	99US-0138847P.
				PR	14-JUN-1999;	99US-0139119P.
Qy	312	lAlaSerSerSerLys-ArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrA	332	PR	16-JUN-1999;	99US-0139452P.
				PR	16-JUN-1999;	99US-0139453P.
Db	898	CTCATCAACCACTAGCAAGAACTACGATTGGTTTG--CACATCTCAAGAAAGAGATATT	955	PR	17-JUN-1999;	99US-0139492P.
				PR	18-JUN-1999;	99US-0139454P.
Qy	332	rgThrProLeuLeuGlyIleGlyLeuLeuVal	343	PR	18-JUN-1999;	99US-0139455P.
		:::		PR	18-JUN-1999;	99US-0139456P.
Db	956	GGTCTCCATTGATGGTTGGAATTGCTCACTTATG	990	PR	18-JUN-1999;	99US-0139457P.
				PR	18-JUN-1999;	99US-0139458P.
RESULT 7						
AAC39099						
ID	AAC39099 standard; DNA; 1683 BP.					
XX						
AC	AAC39099;					
XX						
DT	17-OCT-2000 (first entry)					
XX						
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 23364.					
XX						
KW	Hybridisation assay; genetic mapping; gene expression control;					
KW	protein identification; signal transduction pathway; metabolic pathway;					
KW	promoter; termination sequence; ss.					
XX						
OS	Arabidopsis thaliana.					
XX						
PN	EP1033405-A2.					
XX						
PD	06-SEP-2000.					
XX						
PF	25-FEB-2000; 2000EP-00301439.					
XX						
PR	25-FEB-1999; 99US-0121825P.					
PR	05-MAR-1999; 99US-0123180P.					
PR	09-MAR-1999; 99US-0123548P.					
PR	23-MAR-1999; 99US-0125788P.					
PR	25-MAR-1999; 99US-0126264P.					
PR	29-MAR-1999; 99US-0126785P.					
PR	01-APR-1999; 99US-0127462P.					
PR	06-APR-1999; 99US-0128234P.					
PR	08-APR-1999; 99US-0128714P.					
PR	16-APR-1999; 99US-0129845P.					
				PR	19-APR-1999;	99US-0130077P.
				PR	21-APR-1999;	99US-0130449P.
				PR	23-APR-1999;	99US-0130510P.
				PR	23-APR-1999;	99US-0130891P.
				PR	28-APR-1999;	99US-0131449P.
				PR	30-APR-1999;	99US-0132048P.
				PR	30-APR-1999;	99US-0132407P.
				PR	04-MAY-1999;	99US-0132484P.
				PR	05-MAY-1999;	99US-0132485P.
				PR	06-MAY-1999;	99US-0132486P.
				PR	06-MAY-1999;	99US-0132487P.
				PR	07-MAY-1999;	99US-0132863P.
				PR	11-MAY-1999;	99US-0134256P.
				PR	14-MAY-1999;	99US-0134218P.
				PR	14-MAY-1999;	99US-0134219P.
				PR	14-MAY-1999;	99US-0134221P.
				PR	14-MAY-1999;	99US-0134370P.
				PR	18-MAY-1999;	99US-0134768P.
				PR	19-MAY-1999;	99US-0134941P.
				PR	20-MAY-1999;	99US-0135124P.
				PR	21-MAY-1999;	99US-0135353P.
				PR	24-MAY-1999;	99US-0135629P.
				PR	25-MAY-1999;	99US-0136021P.
				PR	27-MAY-1999;	99US-0136392P.
				PR	28-MAY-1999;	99US-0136782P.
				PR	01-JUN-1999;	99US-0137222P.
				PR	03-JUN-1999;	99US-0137528P.
				PR	04-JUN-1999;	99US-0137502P.
				PR	07-JUN-1999;	99US-0137724P.
				PR	08-JUN-1999;	99US-0138094P.
				PR	10-JUN-1999;	99US-0138540P.
				PR	10-JUN-1999;	99US-0138847P.
				PR	14-JUN-1999;	99US-0139119P.
				PR	16-JUN-1999;	99US-0139452P.
				PR	16-JUN-1999;	99US-0139453P.
				PR	17-JUN-1999;	99US-0139492P.
				PR	18-JUN-1999;	99US-0139454P.
				PR	18-JUN-1999;	99US-0139455P.
				PR	18-JUN-1999;	99US-0139456P.
				PR	18-JUN-1999;	99US-0139457P.
				PR	18-JUN-1999;	99US-0139458P.
				PR	18-JUN-1999;	99US-0139459P.
				PR	18-JUN-1999;	99US-0139460P.
				PR	18-JUN-1999;	99US-0139461P.
				PR	18-JUN-1999;	99US-0139462P.
				PR	18-JUN-1999;	99US-0139463P.
				PR	18-JUN-1999;	99US-0139750P.
				PR	18-JUN-1999;	99US-0139763P.
				PR	21-JUN-1999;	99US-0139817P.
				PR	22-JUN-1999;	99US-0139899P.
				PR	23-JUN-1999;	99US-0140353P.
				PR	23-JUN-1999;	99US-0140354P.
				PR	24-JUN-1999;	99US-0140695P.
				PR	28-JUN-1999;	99US-0140823P.
				PR	29-JUN-1999;	99US-0140991P.
				PR	30-JUN-1999;	99US-0141287P.
				PR	01-JUL-1999;	99US-0141842P.
				PR	01-JUL-1999;	99US-0142154P.
				PR	02-JUL-1999;	99US-0142055P.
				PR	06-JUL-1999;	99US-0142390P.
				PR	08-JUL-1999;	99US-0142803P.
				PR	09-JUL-1999;	99US-0142920P.
				PR	12-JUL-1999;	99US-0142977P.
				PR	13-JUL-1999;	99US-0143542P.
				PR	14-JUL-1999;	99US-0143624P.
				PR	15-JUL-1999;	99US-0144005P.
				PR	16-JUL-1999;	99US-0144085P.
				PR	16-JUL-1999;	99US-0144086P.
				PR	19-JUL-1999;	99US-0144325P.
				PR	19-JUL-1999;	99US-0144331P.
				PR	19-JUL-1999;	99US-0144332P.
				PR	19-JUL-1999;	99US-0144333P.
				PR	19-JUL-1999;	99US-0144334P.





Db	747	AAAGTGGGGCGGTACGAGGTTTGAAGCTGCACTGAGGAAGCTCCGTGGGAAGAAGGCT	806	PR	05-MAR-1999;	99US-0123180P.
Qy	300	AspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThr	319	PR	09-MAR-1999;	99US-0123548P.
Db	807	GATATTTCGGAGGAGGCAGAGATCCAGGATTATATCGAACTCTGGAAGGCTACCG	866	PR	23-MAR-1999;	99US-0125788P.
Qy	320	ThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIle	339	PR	25-MAR-1999;	99US-0126264P.
Db	867	AAAGCCAAGATGCTGGATTGTTTCAGAGGAGATACATACGCTCTGTTCTATAGCTTTC	926	PR	29-MAR-1999;	99US-0126785P.
Qy	340	GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer	359	PR	01-APR-1999;	99US-0127462P.
Db	927	GGGTGATGGTGTTCACAGTTTGGAGGAATCAACGGAATATGTTTCTACACAAGCTCG	986	PR	06-APR-1999;	99US-0128234P.
Qy	360	IlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle	379	PR	08-APR-1999;	99US-0128714P.
Db	987	ATATTGAGCAAGCAGGTTTCCC---ACAAGACTGGGATGATAATATATGCTGTTCTT	1043	PR	16-APR-1999;	99US-0129845P.
Qy	380	GlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeu	399	PR	19-APR-1999;	99US-0130077P.
Db	1044	CAGGTGTAATCACTGCGCTTAATGCACCGATAGTTGACAGAGCCGGAAGAAACCATTG	1103	PR	21-APR-1999;	99US-0130449P.
Qy	400	LeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePhe	419	PR	23-APR-1999;	99US-0130510P.
Db	1104	CTACTGGTTTCTGCAACAGGGTTAGTGATAGGCTGTTTGATAGCAGCGGTTTCTTCTAT	1163	PR	23-APR-1999;	99US-0130891P.
Qy	420	LeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeu	439	PR	28-APR-1999;	99US-0131449P.
Db	1164	CTCAAGGTTTCA-----GACATGGCGCACGAAGCAGTCCCAGTCCTGGCTGTT	1211	PR	30-APR-1999;	99US-0132048P.
Qy	440	ValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeu	459	PR	30-APR-1999;	99US-0132407P.
Db	1212	GTTGGTATAATGGTGTACATAGGATCGTTTTCAGCAGGAATGGGAGCAATGCCGTGGTG	1271	PR	04-MAY-1999;	99US-0132484P.
Qy	460	MetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeu	479	PR	05-MAY-1999;	99US-0132485P.
Db	1272	GTCACTGCTGAGATATTCCCATAAACATAAAGGAGTAGCAGGAGGCATGGCGACGCTG	1331	PR	06-MAY-1999;	99US-0132486P.
Qy	480	AlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSer	499	PR	07-MAY-1999;	99US-0134256P.
Db	1332	GTGAAGTGGTTTGGAGCGTGGGCTGTTTCTTACACTTTCAACTTTCCTCATGTCCTGGAGC	1391	PR	11-MAY-1999;	99US-0134218P.
Qy	500	ValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValVal	519	PR	14-MAY-1999;	99US-0134219P.
Db	1392	TCTTACGGAACTTTCTCTCATTTACGCTGCCATCAACGCGCTGGCCATTGTCTTCGTCATT	1451	PR	14-MAY-1999;	99US-0134370P.
Qy	520	LeuTrpValProGluThrLysGly	527	PR	18-MAY-1999;	99US-0134768P.
Db	1452	GCCATCGTCCTGAGACAAAAGGG	1475	PR	28-MAY-1999;	99US-0136782P.
RESULT 8						
AAC45857						
ID	AAC45857 standard; DNA; 1395 BP.					
XX						
AC	AAC45857;					
XX						
DT	18-OCT-2000 (first entry)					
XX						
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 48015.					
XX						
KW	Hybridisation assay; genetic mapping; gene expression control;					
KW	protein identification; signal transduction pathway; metabolic pathway;					
KW	promoter; termination sequence; ss.					
XX						
OS	Arabidopsis thaliana.					
XX						
PN	EPI033405-A2.					
XX						
PD	06-SEP-2000.					
XX						
PF	25-FEB-2000; 2000EP-00301439.					
XX						
PR	25-FEB-1999; 99US-0121825P.					
PR						

PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.

PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 9.31e-75 Length: 1395  
Score: 933.00 Matches: 180  
Percent Similarity: 65.73% Conservative: 102  
Best Local Similarity: 41.96% Mismatches: 141  
Query Match: 33.57% Indels: 6  
DB: 3 Gaps: 3

US-10-051-909-36 (1-553) x AAC45857 (1-1395)

QY 100 PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe 119  
Db 85 TATCTCAGCACATTTCGTGGGGTCTGTGGTTCTTTTGGTTCGGTTTCGGTTCGGGGATAC 144  
QY 120 SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer 139  
Db 145 TC TTCACCTGCTCAGGCTGCAATTAGGAATGACCTTTCATTGACTATAGCTGAGTTTTC 204  
QY 140 AlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet 159  
Db 205 CTC TTGGTCTCTTACTAACTTTTGGCGCAATGATCGGTGCTATACAAGCGGGCCTATA 264  
QY 160 AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179  
Db 265 GCTGATTAGTTGGAAGAAAGGGCGATGAGAGTTTCCTCTGCATTTTGTGTAGTCGGG 324  
QY 180 TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu 199  
Db 325 TGGCTAGCAATCATCTTTGCCAAGGGAGTAGTGGCTCTGGACCTTGGGAAGACTGGCAACG 384  
QY 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219  
Db 385 GGATATGGAATGGGACCATTTTCCTATGTGGTGCCAAATCTTTATAGCAGAAATTCACCT 444  
QY 220 GlnAsnMetArgGlyAlaLeuGlySerValAsn---GlnLeuSerValThrPheGlyIle 238  
Db 445 AAAACTTTCAGAGGGGCTCTAACCCACACTGAACCAAGCAGATTCCTGATCTGCACCTGGAGTG 504  
QY 239 PheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAla 258

Db 505 TCGGTTTCCTTCATCATTAGGCACACTAGTAGCGTGGAGAGTCTTGGCATTAAATAGGAATC 564  
:::|||||:::|  
Qy 259 LeuProCysThrMetLeuIleProGlyLeuPhePheIleProGlySerProArgTrpLeu 278  
:::|||||  
Db 565 ATCCCATGCGTGCCTCCTTCCTTGGCCCTCTTTTATCCCTGAGTCTCCAAGATGGCTG 624  
:::|||||  
Qy 279 AlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu 298  
|||||:::|  
Db 625 GCAAAAGTGGGCGTGATACGGAGTTTGAAGCTGCACTGAGGAAGCTCCGTGGGAAGAAG 684  
:::|||||  
Qy 299 ThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArg 318  
|||||:::|  
Db 685 GCTGATATTCGGAGGAGGCAGCAGAGATCCAGGATTATATCGAAACTCTGGAAGGCTA 744  
:::|||||  
Qy 319 ThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGly 338  
:::|||||  
Db 745 CCGAAAGCCCAAGATGCTGGATTGTTCAGAGGAGATACATACGCTCTGTTCTTATAGCT 804  
:::|||||  
Qy 339 IleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSer 358  
|||||:::|  
Db 805 TTCGGGTTGATGGTGTTCAGCAGTTTGGAGGAATCAACGGAATATGTTTCTACACAAGC 864  
:::|||||  
Qy 359 SerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAla 378  
|||||:::|  
Db 865 TCGATATTTGAGCAAGCAGGTTTTTCCC--ACAAGACTTGGGATGATAATATATGCTGTT 921  
:::|||||  
Qy 379 IleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIle 398  
:::|||||  
Db 922 CTTCAGGTGTAATCACTGCGCTTAATGCACCGATAGTTGACAGAGCCGGAAGAAACCA 981  
:::|||||  
Qy 399 LeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhe 418  
|||||:::|  
Db 982 TTGCTACTGGTTTCTGCAACAGGGTTAGTGATAGGCTGTTTGATAGCAGCGGTTCTTTC 1041  
:::|||||  
Qy 419 PheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSer 438  
:::|||||  
Db 1042 TATCTCAAGTTTCAC-----GACATGGCGCACGAAGCAGTCCCAGTCTCTGGCT 1089  
:::|||||  
Qy 439 LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp 458  
:::|||||  
Db 1090 GTTGTGGTATAAATGGTGTACATAGGATCGTTCAGCAGGAATGGAGCAATGCCGTGG 1149  
:::|||||  
Qy 459 LeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlySerIleAlaThr 478  
:::|||||  
Db 1150 GTGGTCATGCTGAGATATTTCCCATAAACATAAAAGGAGTAGCAGGAGGCATGGCGACG 1209  
:::|||||  
Qy 479 LeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrp 498  
||| |||||  
Db 1210 CTGGTGAAGTGGTTTGGAGCGTGGGCTGTTCTTACACTTTCAACTTCCTCATGTCCTGG 1269  
||| |||||  
Qy 499 SerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheVal 518  
||| |||||  
Db 1270 AGCTCTTACGGAACCTTTCCTCATTTACGTGCCATCAACGCGCTGGCCATTGCTTCGTC 1329  
||| |||||  
Qy 519 ValLeuTrpValProGluThrLysGly 527  
::: |||||  
Db 1330 ATTGCCATCGTGCCTGAGACAAAGGG 1356  
||| |||||

RESULT 9

AAC45298

ID AAC45298 standard; DNA; 1685 BP.

XX

AC AAC45298;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 46007.

XX

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway; metabolic pathway;

KW promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.



PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145113P.  
PR 22-JUL-1999; 99US-0145114P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.

PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 9.97e-74  
Score: 923.00  
Percent Similarity: 65.58%  
Best Local Similarity: 42.09%  
Query Match: 33.21%  
DB: 3

Length: 1685  
Matches: 181  
Conservative: 101  
Mismatches: 141  
Indels: 7  
Gaps: 3

US-10-051-909-36 (1-553) x AAC45298 (1-1685)

Qy 100 PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe 119  
Db 207 TATCTCAGCACATTCGTTGCGGTCCTGTGGTCTTTTTCGGTTCGGTCTTGTGCGGGATAC 266  
Qy 120 SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer 139  
Db 267 TCTTCACCTGCTCAGGCTGCAATTAGGAATGACCTTTTCATTGACTATAGCTGAGTTTCA 326  
Qy 140 AlapheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet 159  
Db 327 CTCCTTTGGTTCCTTACTAACTTTTGGCGCAATGATCGGTGCTATACAAGCGGGCTATA 386  
Qy 160 AlaGluTyriIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly 179  
Db 387 GCTGATTTAGTTGGAAGAAAGGGCGCATGAGAGTTTCTCTGCAATTTTGTGAGTCGGG 446  
Qy 180 TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu 199  
Db 447 TGGCTAGCAATCATCTTTTGCCACAGGAGTAGTGGCTCTGGACCTTGGAGACTGGCAACG 506

Qy 200 GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro 219  
Db 507 GGATATGGAATGGGAGCATTTTCCTATGTGGTGCCAATCTTTATAGCAGAAATTGCACCT 566  
Qy 220 GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu---Ser-ValThrPheGlyI1 238  
Db 567 AAAACTTTCAGAGGGGCTCTAACACACTGAACACAGACAGATTCTGTATTCGCACTGGAGT 626  
Qy 238 ePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAl 258  
Db 627 GTCGGTTTCCTTCATCATAGGCACACTAGTGAGCTGGAGAGTCTTGGCAATTAATAGGAAT 686  
Qy 258 aLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLe 278  
Db 687 CATCCCATGCGCTGCTCCTTCCTTGGCCTCTTTTATCCCTGAGTCTCCAAGATGGCT 746  
Qy 278 uAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheG1 298  
Db 747 GGCAAAAGTGGGGCGTGATACGGAGTTTGAAGCTGCACTGAGGAAGCTCCGTGGGAAGAA 806  
Qy 298 uThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysAr 318  
Db 807 GGCTGATATTTCCGAGGAGGCAGCAGAGATCCAGGATTATATCGAAACTCTGGAAGGCT 866  
Qy 318 gThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlu 338  
Db 867 ACCGAAAGCCAGATGCTGGATTGTGTTTCAGAGGAGATACATACGCTCTGTCTTATAGC 926  
Qy 338 yIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlase 358  
Db 927 TTTCCGGTTGATGGTGTGTTTCAGCAGTTTGGAGGAATCAACGGAATATGTCTTACACAAG 986  
Qy 358 rSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAl 378  
Db 987 CTCGATATTTGAGCAAGCAGGTTTTCCTCC---ACAAGACTTGGGATGATAATATATGCTGT 1043  
Qy 378 aileGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgI1 398  
Db 1044 TCTTCAGGTGGTAATCACTCGCTTAATGCACCGATAGTTGACAGAGCCGGAAGAAACC 1103  
Qy 398 eLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPh 418  
Db 1104 ATTGCTACTGGTTCTGCAACAGGGTTAGTGATAGGCTGTTTGATAGCAGCGGTTTCTTT 1163  
Qy 418 ePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSe 438  
Db 1164 CTATCTCAAGTTTCAC-----GACATGGCGCACGAAGCAGTCCAGTCTCTGGC 1211  
Qy 438 rLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTr 458  
Db 1212 TGTTGTTGGTATAATGGTGATAGGATCGTTTTCAGCAGGAATGGGAGCAATGCCGTG 1271  
Qy 458 pLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaTh 478  
Db 1272 GGTGTCATGTCGTAGATATTTCCCATATAACATAAAGGAGTAGCAGGAGGCATGGCGAC 1331  
Qy 478 rLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTr 498  
Db 1332 GCTGGTGAACCTGGTTTGGAGCGTGGGCTGTTTCTTACACTTCAACTTCCTCATGTCTG 1391  
Qy 498 pSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheVa 518  
Db 1392 GAGCTCTTACGGAACCTTTCCTCATTTTACGTGCGCATCAACGCGCTGGCCATGTCTTCGT 1451  
Qy 518 lValLeuTrpValProGluThrLysGly 527  
Db 1452 CATTGCCATCGTGCCTGAGACAAAAGGG 1479  
RESULT 10  
AAC45290  
ID AAC45290 standard; DNA; 1743 BP.  
XX  
AC AAC45290;

XX 18-OCT-2000 (first entry)  
DT  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 45977.  
DE  
XX Hybridisation assay; Genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX Arabidopsis thaliana.  
OS  
XX EP1033405-A2.  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.







138	Qy	PhSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGly	157
		:::     :::     :::     :::     :::     :::     :::     :::	
271	Db	TACTCAATGTTCCGGTCGATCTTGACATTAGGAGGCTTGATCGGTGCAGTATTACGCGGT	330
		:::     :::     :::     :::     :::     :::     :::     :::	
158	Qy	GlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIle	177
		:::     :::     :::     :::     :::     :::     :::     :::	
331	Db	AAAGTCGCTGATGTTCTGGGAAGAAACCGGACGATGTTGTTTTCGGAATCTTCTGTATC	390
		:::     :::     :::     :::     :::     :::     :::     :::	
178	Qy	IleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeu	197
		:::     :::     :::     :::     :::     :::     :::	
391	Db	ACAGCTGGCTTTGTGTAGCATTTGGCTCAGAAATGCAATGTGGCTGGACTGTGGAAGATTG	450
		:::     :::     :::     :::     :::     :::     :::	
198	Qy	LeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIle	217
		:::     :::     :::     :::     :::     :::     :::	
451	Db	TTACTTGAATCCGGCTGGTATATTAGCTACGTGATTCGGGTGTATATAGCCGAAATT	510
		:::     :::     :::     :::     :::     :::     :::	
218	Qy	SerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGly	237
		:::     :::     :::     :::     :::     :::     :::     :::	
511	Db	GCACCTAAACATGTCGGAGGATCGTTTGTGTTTCGCCAATCAGTTGATGCAAAATTGCGGA	570
		:::     :::     :::     :::     :::     :::     :::	
238	Qy	IlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGly	257
		:::     :::     :::     :::     :::     :::     :::	
571	Db	ATTTCACTCTTCTTCATCATTTGGCAATTTTATTCATGGAGACTACTAACAGTAGTCGA	630
		:::     :::     :::     :::     :::     :::     :::	
258	Qy	AlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrp	277
		:::     :::     :::     :::     :::     :::     :::     :::	
631	Db	TTGGTGCATGTGTGTTCTCACGCTCTTTTGTATTTTTCATCCCCGAATCTCCAAGATGG	690
		:::     :::     :::     :::     :::     :::     :::	
278	Qy	LeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPhe	297
		:::     :::     :::     :::     :::     :::     :::	
691	Db	CTGGCGAAGTTAGGTCGTGATAAAGAAATGCCGATCTTCGTTGCAACGCCCTTAGGGATCT	750
		:::     :::     :::     :::     :::     :::     :::	
298	Qy	GluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLys	317
		:::     :::     :::     :::     :::     :::     :::     :::	
751	Db	GACGTCGATATTCTCGTGAAGCAACACACAATTAGAGATACCATTGACATGACAGAAAAAC	810
		:::     :::     :::     :::     :::     :::     :::	
318	Qy	ArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeu	337
		:::     :::     :::     :::     :::     :::     :::	
811	Db	GGTGTGAAACTAAGATGCTGAATTGTTTCAGAGACGATACGCATATCCGTTAATTATC	870
		:::     :::     :::     :::     :::     :::     :::	
338	Qy	GlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAla	357
		:::     :::     :::     :::     :::     :::     :::	
871	Db	GGAGTTGGTTTAAATGTTTGTGCAACAATTGTGTGGGAGCTCCGGTGTACCTATTATGCT	930
		:::     :::     :::     :::     :::     :::     :::	
358	Qy	SerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGly	377
		:::     :::     :::     :::     :::     :::     :::	
931	Db	AGTAGCCTCTTCAACAAAGGAGGATTTCCA---AGTGCTATTGGCACATCCGTAATAGCC	987
		:::     :::     :::     :::     :::     :::     :::	
378	Qy	AlaIleGlnValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArg	397
		:::     :::     :::     :::     :::     :::     :::	
988	Db	ACAATTATGGTTCCTCAAAAGCAATGCTGGCAACACAGCTCTAGTCGATAAAATGGGAGGAGA	1047
		:::     :::     :::     :::     :::     :::     :::	
398	Qy	IleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValVal	417
		:::     :::     :::     :::     :::     :::     :::	
1048	Db	ACGCTCCTAATGGCTCTTGTCTTGCAATGGGTTTGAGTGCTTTGCTCTTAAAGTGT---	1104
		:::     :::     :::     :::     :::     :::     :::	
418	Qy	PhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeu-----	434
		:::     :::     :::     :::     :::     :::     :::	
1105	Db	-----TCTTACGGTTTTCCAGTCGTTTGGCATTTCTCCAGAACTC	1143
		:::     :::     :::     :::     :::     :::     :::	
435	Qy	---ThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMet	453
		:::     :::     :::     :::     :::     :::     :::     :::	
1144	Db	ACTCCCATCTTCACTTGCATCGGCGTCTTGGGTACACATGTGTTCATTTGCCATGGGAATG	1203
		:::     :::     :::     :::     :::     :::     :::	
454	Qy	GlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGly	473
		:::     :::     :::     :::     :::     :::     :::	
1204	Db	GGAGGACTACCATGGATATATATGGCTGAGATATTTCCGATGAATGTGAAAGTGTGCAGCT	1263
		:::     :::     :::     :::     :::     :::     :::	
474	Qy	GlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsn	493
		:::     :::     :::     :::     :::     :::     :::	
1264	Db	GGGACCTTAGTTACTGTAAACCAATTGGTTATTGTTGGTGGATTATCACATACACTTCAAT	1323
		:::     :::     ::	

Db	1324	TTTATGCTAGAAATGGAATGCATCAGGAATGTTCTCTCATCTTCTCAATGGTCTCCGCCAGT	1383
Qy	514	ThrIleValPheValValLeuTrpValProGluThrLysGly***AsnSerArgGlyAsp	533
Db	1384	TCGATCGTATTTATATATACTTTTGGTACCTGAGACAAAAGGCCGATC	1443
Qy	534	ThrIlePheValSerLeuSerIleGln	542
Db	1444	CAAGCACTGCTCAACAACTCTGTGCAA	1470
RESULT 12			
AAC40459			
ID	AAC40459 standard; DNA; 1755 BP.		
AC	AAC40459;		
XX			
XX	17-OCT-2000 (first entry)		
DT			
XX	Arabidopsis thaliana DNA fragment SEQ ID NO: 28375.		
DE			
XX			
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway; metabolic pathway;		
KW	promoter; termination sequence; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-00301439.		
XX			
PR	25-FEB-1999; 99US-0121825P.		
PR	05-MAR-1999; 99US-0123180P.		
PR	09-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	06-MAY-1999; 99US-0132487P.		
PR	07-MAY-1999; 99US-0132863P.		
PR	11-MAY-1999; 99US-0134256P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		
PR	18-MAY-1999; 99US-0134768P.		
PR	19-MAY-1999; 99US-0134941P.		
PR	20-MAY-1999; 99US-0135124P.		
PR	21-MAY-1999; 99US-0135353P.		
PR	24-MAY-1999; 99US-0135629P.		
PR	25-MAY-1999; 99US-0136021P.		
PR	27-MAY-1999; 99US-0136392P.		
PR	28-MAY-1999; 99US-0136782P.		
PR	01-JUN-1999; 99US-0137232P.		
PR	03-JUN-1999; 99US-0137528P.		
PR	04-JUN-1999; 99US-0137502P.		
PR	07-JUN-1999; 99US-0137724P.		
PR	08-JUN-1999; 99US-0138094P.		

PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores: Length: 1755  
Pred. No.: 4.59e-73 Matches: 191  
Score: 916.00 Conservative: 96  
Percent Similarity: 61.99% Mismatches: 165  
Best Local Similarity: 41.25%



Query Match: 32.96% Indels: 12  
DB: 3 Gaps: 3

US-10-051-909-36 (1-553) x AAC40459 (1-1755)

QY	94	Glus	His	Val	Ser	Ala	Phe	Leu	Cys	Thr	Leu	Ile	Val	Ala	Leu	Gly	Pro	Ile	Gln	Phe	113
Db	188	GAAT	CAT	ATTT	TG	TG	TT	CT	CT	TT	CAG	TAC	TTT	TG	TG	CT	GT	GT	GG	CT	247
QY	114	Gly	Phe	Thr	Ser	Gly	Phe	Ser	Ser	Pro	Thr	Gln	Asp	Ala	Met	Val	Arg	Asp	Leu	Asn	133
Db	248	GGCT	CTT	GT	TG	GAT	ACT	CAG	CGC	CT	ACT	CAG	TCA	TCT	ATA	AG	ACA	AG	AT	CT	307
QY	134	Ser	Ile	Ser	Glu	Phe	Ser	Ala	Phe	Gly	Ser	Leu	Ser	Asn	Val	Gly	Gly	Met	Val	Gly	153
Db	308	TCC	CTT	GC	AG	AGT	TC	CC	AT	GT	TT	TG	GA	TCC	AT	CT	TAA	CT	AT	CG	367
QY	154	Ile	Ala	Ser	Gly	Gln	Met	Ala	Glu	Tyr	Ile	Gly	Arg	Lys	Gly	Ser	Leu	Met	Ile	Ala	173
Db	368	GTT	AT	GAG	TGG	AAA	ATTT	CAG	ATT	TCT	CCG	GCC	GAAA	AGG	GG	CA	AT	TAG	GAC	GT	427
QY	174	Ile	Pro	Asn	Ile	Ile	Gly	Trp	Leu	Ala	Ile	Ser	Phe	Ala	Lys	Asp	Ala	Ser	Phe	Leu	193
Db	428	TGCT	TT	CG	CA	TT	AC	AG	TT	GG	CT	CG	CT	CT	CT	TC	CA	AG	GG	GG	487
QY	194	Met	Gly	Arg	Leu	Glu	Gly	Phe	Gly	Val	Gly	Ile	Ile	Ser	Tyr	Thr	Val	Pro	Val	Tyr	213
Db	488	GTAG	GA	AG	GT	CT	TT	TA	CAG	GAT	TG	GA	TTT	TCT	TAT	GT	GG	TCC	CT	GT	547
QY	214	Ile	Ala	Glu	Ile	Ser	Pro	Gln	Asn	Met	Arg	Gly	Ala	Leu	Gly	Ser	Val	Asn	Gln	Leu	233
Db	548	ATT	GT	CT	GAG	AT	AT	CT	CC	AA	GA	AT	CT	CC	GAG	TG	GACT	CA	CA	ACT	607
QY	234	Val	Thr	Phe	Gly	Ile	Phe	Leu	Ala	Tyr	Leu	Leu	Gly	Met	Phe	Ile	Pro	Trp	Arg	Leu	253
Db	608	ANT	GT	GAT	CGG	CT	CAT	CGG	TTT	CTT	CT	TG	AT	CGG	AT	CT	CT	CAT	TTT	CT	667
QY	254	Ala	Val	Ile	Gly	Ala	Leu	Pro	Cys	Thr	Met	Leu	Ile	Pro	Gly	Leu	Phe	Phe	Ile	Pro	273
Db	668	GCC	TA	ACC	GTA	CTT	TG	CT	CC	TG	CAT	TG	TTT	TG	CT	CT	TG	GG	CT	TG	727
QY	274	Ser	Pro	Arg	Trp	Leu	Ala	Lys	Met	Asn	Leu	Thr	Glu	Asp	Cys	Glu	Thr	Ser	Leu	Gln	293
Db	728	TCT	CT	CG	AT	GG	CT	GG	CA	AA	AG	CA	GG	CC	AT	CA	AG	AA	GAG	AT	787
QY	294	Leu	Arg	Gly	Phe	Glu	Thr	Asp	Ile	Thr	Thr	Glu	Val	Asn	Asp	Ile	Lys	Arg	Ala	Val	313
Db	788	CTG	CAG	GAA	AA	GAT	GC	AG	AT	AT	CA	AA	AT	GA	GC	AG	CG	GT	AT	CA	847
QY	314	Ser	Ser	Ser	Lys	Arg	Thr	Thr	Ile	Ser	Phe	Gln	Glu	Leu	Asn	Gln	Lys	Lys	Tyr	Arg	333
Db	848	GCT	CT	AG	AG	AT	CT	TC	CA	AA	AG	CA	AG	AT	CC	AG	AC	CT	TG	TG	907
QY	334	Pro	Leu	Leu	Gly	Ile	Gly	Leu	Leu	Val	Leu	Gln	Asn	Leu	Ser	Gly	Ile	Asn	Gly	Val	353
Db	908	TCT	GT	CAT	CAT	TG	TG	TT	CC	CT	GAT	GG	TAT	TC	CA	AC	AG	TTT	TG	TG	967
QY	354	Leu	Phe	Tyr	Ala	Ser	Ser	Ile	Phe	Lys	Ala	Ala	Gly	Val	Thr	Asn	Ser	Asp	Leu	Ala	373
Db	968	GG	AT	CT	AC	GCA	AG	TG	AA	AC	GT	TT	TG	TAA	AG	CC	GG	AT	TAC	CT	1027
QY	374	Cys	Ser	Leu	Gly	Ala	Ile	Gln	Val	Leu	Ala	Thr	Gly	Val	Thr	Thr	Trp	Leu	Asp	Arg	393
Db	1028	A	-	TCG	CT	AT	CG	CT	TG	TT	CA	GT	GCC	AA	TAA	CT	GT	TT	GG	AA	1085
QY	394	Ala	Gly	Arg	Arg	Ile	Leu	Leu	Ile	Ile	Ser	Thr	Ser	Gly	Met	Thr	Leu	Cys	Leu	Ala	413
Db	1086	TCT	GG	AC	GA	AG	GC	CA	CT	AA	TAT	GAT	TT	CA	GT	GG	TAT	CT	TCT	TG	1145
QY	414	Val	Ser	Val	Val	Phe	Phe	Leu	Lys	Asp	Asn	Ile	Ser	Gln	Asp	Ser	Asn	Ser	Tyr	Tyr	433
Db	1146	AC	AG	GC	AC	AT	CT	TT	CT	TCT	ACT	CA	G	-	-	-	-	-	-	-	1187

CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office  
XX  
SQ Sequence 1359 BP; 340 A; 238 C; 327 G; 454 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 1..11e-71 Length: 1359  
Score: 899.00 Matches: 188  
Percent Similarity: 60.17% Conservative: 99  
Best Local Similarity: 39.41% Mismatches: 149  
Query Match: 32.35% Indels: 42  
DB: 6 Gaps: 5  
US-10-051-909-36 (1-553) x ABZ13822 (1-1359)  
QY 79 GlnSerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer 98  
Db 19 GAAGCTAATCTTGCTCCTGAAACTTCA-----CTTATTAAACAAGAGAATCAAGACTCA 72  
QY 99 Ala-----PheLeuCysThrLeuIleValAlaLeuGlyProIle 111  
Db 73 TCGGCGACAATAACAACACTACTCTCTTGCTCACCACCTTTGTAGCCGTGTCGTTCCTTT 132  
QY 112 GlnPheGlyPheThrSerGlyPheSerProThrGlnAspAlaMetValArgAspLeu 131  
Db 133 GTCCTTGGCTCAGCTATAGGATATTCACTCACCTGTTCAGTCGGATTTAACCAAGAGTTG 192  
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151  
Db 193 AATCTCTGTGGCAGAGTACTCGCTCTTCGGTTTCGATTCTTGACAATTGGAGCAATGATC 252  
QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171  
Db 253 GCGCAGCTATGAGTGGACGAATCGCGATATGATAGGCAGCAGCTACTATGGGTTTC 312  
QY 172 AlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAlaSerPhe 191  
Db 313 TCGGAGATGTTTTCATTTCTCGGCTGGTTAGCAATCTACCTTTCTAAGTTGCAATTTGG 372  
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211  
Db 373 CTTGACGTTGGGAGATTCTTGGTTGATTGGAATGGGCGTTTTTTCGTTGTGGTTCCCT 432  
QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231  
Db 433 GTATACATAGCTGAATCACTCCTTAAGGCTCTCGTGGTGGTTTACAACTGTTTCATCAG 492  
QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251  
Db 493 TTA CTGATATGTTGGGTGCTCAGTCACATATCTCTGGGCTCTTTTATAGGCTGGCGA 552  
QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271  
Db 553 ATTTTAGCTTTGATCGGAATGATACCATGTGTGTACAAATGATGGGATTATTGTTCATC 612  
QY 272 ProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291  
Db 613 CCAGAGTCTCCTAGATGGTTGGCAAAAGTTGGCAAAATGGGAAGAAATTTGAGATTGCATTG 672  
QY 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311  
Db 673 CAACGATTACGTGGTGAATCCGCAGATATTTCATATGAGTCTAATGAGATCAAAAGATTAT 732  
QY 312 ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331  
Db 733 ACGCGAAGACTTACAGATTATTACAGAAGGTACTATAGTATCTGTTCCAGCCACAGTAT 792  
QY 332 ArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351  
Db 793 GCTAAATCTCTAGTTGTGGGAGTTGGTCTGTGAGTTTTCACAACAATTTGGAGGGGTTAAT 852  
QY 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371

Db 853 GGAATTGCTTTTACGCTAGCTCTATTTTTCGAATCTGCTGGGTT---TCTTCCAAAAATT 909  
QY 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLeuLeu 391  
Db 910 GGAATGATTGCGATGGTGGTCGTACAGATCCCTATGACAACTTTAGGTGTTACTGTTGATG 969  
QY 392 AspArgAlaGlyArgArgIleLeuLeuIleSerThrSerGlyMetThrLeuCysLeu 411  
Db 970 GATAAATCAGGAAGACGACCAATTACTTCTAATTTCTGCTACTGGAACATGTATCGGGTGT 1029  
QY 412 LeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431  
Db 1030 TTCCTCGTCGGTTTATCATTTTCAATTACAG----- 1059  
QY 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451  
Db 1060 -----GTGTACACCGGATCATTTTCATTA 1083  
QY 452 GlyMetGlyAlaIleProTyrPleuMetMetSerGluIleLeuProValSerIleLysSer 471  
Db 1084 GGGATGGGTGGGATTCCTTGGGTCAATTATGTCCAGAGATATTTCCAATAGACATAAAGGA 1143  
QY 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTyrLeuThrSerPheAlaIleThrMetThr 491  
Db 1144 TCAGCCGGAAGCCTCGTGACTGTTGTAGCTGGGTCGGATCATGGATTATATCTTTTACA 1203  
QY 492 ThrAsnLeuMetLeuThrTyrSerValGlyGlyThrPheLeuSerTyrMetValValSer 511  
Db 1204 TTTAACCTTTCTAATGAATTGGAATCCGCGCAGGAACGTTTATGTTTTCGCTACTGTTTGT 1263  
QY 512 AlaPheThrIleValPheValValLeuTyrPleuProGluThrLysGly\*\*\*AsnSerArg 531  
Db 1264 GGGGCTACTGTTATTTTGTAGCAAAACTCGTACCAGAAACCAAGTCG-CACACTTGA 1322  
QY 532 GlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeuGlnTyrLeu 548  
Db 1323 GGAATCCAATATTC-----GATTGGTTA 1346  
RESULT 14  
ADM47939  
ID ADM47939 standard; DNA; 1768 BP.  
XX ADM47939;  
XX 03-JUN-2004 (first entry)  
XX Polynucleotide sequence #357 useful in producing transgenic plants.  
XX Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;  
KW osmotic stress; sugar transport; cell cycle pathway; plant height;  
KW carbohydrate transport; crop productivity; plant growth;  
KW stress resistance; disease resistance; insect resistance; heat tolerance;  
KW nitrogen assimilation; water stress tolerance;  
KW photosynthetic carbon fixation; virus resistance; gene therapy; gene; do.  
XX Glycine max.  
OS US2003233670-A1.  
XX 18-DEC-2003.  
XX 04-DEC-2002; 2002US-00310154.  
XX 04-DEC-2001; 2001US-0337358P.  
XX (EDGE/) EDGERTON M D.  
PA (CHOM/) CHOMET P S.  
PA (LACC/) LACCETTI L B.  
XX Edgerton MD, Chomet PS, Laccetti LB;  
XX WPI; 2004-061374/06.  
DR

DR P-PSDB; ADM48307.  
XX New polynucleotide, useful for manipulating plant protein quality.  
PT improving plant growth, yield and crop productivity or grain composition  
PT or producing plants with improved properties.  
XX Claim 1; SEQ ID NO 357; 144pp; English.  
PS  
XX The present invention relates to polynucleotide sequences, and the  
CC proteins they encode. The sequences are isolated from a variety of  
CC organisms such as plants (e.g. maize, rice, sorghum, thale cress,  
CC soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The  
CC polynucleotide and polypeptide sequences of the invention are useful in  
CC the production of transgenic plants that have improved properties. Also  
CC disclosed are methods of producing fertile transgenic plants, preferably  
CC maize, with desired phenotypes. The polynucleotide and polypeptide  
CC sequences are useful for improving plants by providing protection against  
CC osmotic stress, improving altering sugar transport and/or metabolism,  
CC modifying the cell cycle pathway, reducing plant height, modifying  
CC carbohydrate transport, improving crop productivity, improving plant  
CC growth and stress resistance, improving disease resistance, improving  
CC insect resistance, improving cold or heat tolerance, improving nitrogen  
CC assimilation, improving stalk strength, improving water stress tolerance,  
CC improving photosynthetic carbon fixation, improving biotic and abiotic  
CC stress resistance, improving senescence, and conferring virus resistance.  
CC increased vigour, reducing senescence, and conferring virus resistance.  
CC The present sequence represents a polynucleotide sequence of the  
CC invention. Note: The sequence data for this patent is not provided in the  
CC printed specification but is obtained in electronic format from the USPTO  
CC website at seqdata.uspto.gov.  
XX  
SQ Sequence 1768 BP; 516 A; 300 C; 384 G; 568 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 5,73e-71 Length: 1768  
Score: 893.00 Matches: 191  
Percent Similarity: 56.25% Conservative: 97  
Best Local Similarity: 37.30% Mismatches: 173  
Query Match: 32.13% Indels: 51  
DB: 12 Gaps: 4

US-10-051-909-36 (1-553) x ADM47939 (1-1768)  
QY 43 SerAsnArgGly-GlyAlaGlyAlaGlyGluGluSerGlySerAspHisAspGlyValle 62  
Db 119 TCACACCAGAGGATGCTGCTGCTTGTAAAGAAATGGATCAGAT----- 164  
QY 62 uArgArgProLeuLeuAsnThrGlySerTyrArgMetSerSerArgGlnSerSerPh 82  
Db 165 -----AAGAGTGT 172  
QY 82 eAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCy 102  
Db 173 GAAAAATGGATCCATTGGAATGGTTTG-----CTTAG 205  
QY 102 sThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerPr 122  
Db 206 CACACTGTGTGCTGTTGTGGTCTTTCACATTGGAACTTGTGTGGGCTATTTCAGCACC 265  
QY 122 oThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGl 142  
Db 266 CACTCAAGCGGCTATCAGGGCAGATCTTAATCTCTCTGCTGAGTTTCCATGTTTGG 325  
QY 142 ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162  
Db 326 TTCATTAGTGACCATTGGTGCAATGCTTGGAGCTATAACACAGCGGACGAGATTACAGATT 385  
QY 162 rIleGlyArgIysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182  
Db 386 CATTGGCCGCAAGGGCAATGAGGATTTCACACAGGATTTTGCATTACAGGATGGATAGC 445  
QY 182 aIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGl 202

Db 446 TGTCTTCTTCTCTAAGGGTTCTTACTCCCTTGACTTCGGAAGATTTTTTCACAGGTTACGG 505  
QY 202 YValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMe 222  
Db 506 CATTTGGAGTTATCTCATACGTGGTTCTGTATATATAGCAGAAATAGCACCCAAAATCT 565  
QY 222 tArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTy 242  
Db 566 TCGAGGAGGACTAGCAACAACAATCAGCTTTTGATTGTTACTGGAGGGTCAGTCTCAT 625  
QY 242 rLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuProCysTh 262  
Db 626 CTTATTAGGAAGTGTATAAATTGGAGAGAACTTTGCATTAGCGGGGCTAGTGCCTTGCAT 685  
QY 262 rMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrLeuAlaLysMetAs 282  
Db 686 TTGCTTGTGCTGGTGTGTGCTTTATCCCTGAGTCCCGCAGATGGCTGGCTAAGGTTGG 745  
QY 282 nLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleTh 302  
Db 746 CCGTGAAAAAGAAATTCAACTAGCTTTAAAGTAGACTTCGGGGTAAAGATGCTGATATTC 805  
QY 302 rThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSe 322  
Db 806 TGATGAAGCTGCTGAAATCTGGATTATATTGAAACTCTTCAAAGTCTTCTTAAGACTAA 865  
QY 322 rPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuLe 342  
Db 866 GCTGTTGGATTTGTCCAAAGCAAAATATGTGCACTCTGTAGTTATGGGTCGGCTTAAT 925  
QY 342 uValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLy 362  
Db 926 GGCATGTCAACAATCTGTGGAATTAATGGCATAGGATTTCTACAGCTGAGATTTTGT 985  
QY 362 sAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValle 382  
Db 986 AGCAGCTGGACTTTCTTCAGGAAAAGCTGGTACCATAGCATATGCTTGTATACAGATTCC 1045  
QY 382 uAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIle 402  
Db 1046 ATTTACATTATTGGGAGCCATTTTGATGGACAAGTCTGGAAGAAGACCTCTGTAATGGT 1105  
QY 402 eSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAs 422  
Db 1106 TTCCGCGAGCTGGGACATTTCTTGGTTGCTTTGTGTGCTGCCTTTGCTTCTCTCAAGGA 1165  
QY 422 pAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyI 442  
Db 1166 C-----CAAAGCTTATTGCCTGAGTGGGTACCTATATTAGCATTTGCTGGCGT 1213  
QY 442 eValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTyrLeuMetMetSe 462  
Db 1214 TCTGATCTACATAGCAGCATTTTCAATTGGGCTCGGATCAGTTCATGGGTGATAATGTC 1273  
QY 462 rGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTr 482  
Db 1274 TGAGATAATTTCCCATACATTTGAAGGGGACTGCTGGAAGCTTGGTGGTTTGGTGGCTTG 1333  
QY 482 pLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTyrSerValGlyGl 502  
Db 1334 GCTAGGAGCTTGGGTAGTTTCATATATACTTTCAACTTTCTATAGAGCTGGAGTTCTCCTGG 1393  
QY 502 yThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpVa 522  
Db 1394 TACATTGTTTTGTATGCTGGATGTTCCCTCTTAACCTATTCTATTATTGTAGCAAAATTAGT 1453  
QY 522 lProGluThrLysGly\*\*\*AsnSerArgGlyAspThrIlePheValSerLeuSerIleGl 542  
Db 1454 CCCAGAAACCAAGGA----- 1469  
QY 542 nArgGlnLeuGlnTrpLeuProGluCysLeuSer 553  
Db 1470 -AAAACTTTGGAAGAGATCCAGGCTTGTATTAGT 1502



RESULT 15  
ABZ13486  
ID ABZ13486 standard; DNA; 1407 BP.  
XX  
AC ABZ13486;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1291.  
XX  
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 24-AUG-2001; 2001WO-US026685.  
XX  
PR 24-AUG-2000; 2000US-0227866P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

PI Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.

XX Claim 144; SEQ ID NO 1291; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising: (a) contacting nucleic acid  
CC representative of expressed polynucleotides in the plant cell with an  
CC array or probes representative of the plant cell genome; and (b)  
CC detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention. Note: The sequence data for this patent is  
CC not represented in the printed specification but is based on sequence  
CC information supplied to Derwent by the European Patent Office

XX SQ Sequence 1407 BP; 350 A; 261 C; 331 G; 465 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.05e-69 Length: 1407  
Score: 877.50 Matches: 179  
Percent Similarity: 59.06% Conservative: 98  
Best Local Similarity: 38.17% Mismatches: 166  
Query Match: 31.58% Indels: 26  
DB: 6 Gaps: 5

US-10-051-909-36 (1-553) x ABZ13486 (1-1407)

Qy 66 LeuLeuAsnThrGlySerTrpTyrArg-MetSerSerArgGlnSerSerPheAlaProGl 85  
Db 12 CTCAAGAACTTGGAAAGCTGGTTTGCTACTGAGGAAGAACCAACGAC----- 60  
Qy 85 yThrSerSerMetAlaValLeuArgGluSerHisValSerAla-----PheLeuCysTh 103  
Db 61 -----ATCAACGAATGTCGTATCACTGCTGTGTACTTTTCAGTAC 101  
Qy 103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh 123  
Db 102 TTTCGTTTCTGTTGTGGCTCTTTCTGCTCGGTTGTGGCGCAGGTTATTTCATCAGTTGC 161

Qy 123 rGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySe 143  
Db 162 TCAAAACAGGGATCATAAATGATTTAGTCTCTCTGTGTCACAATACTCCATGTTGGTTC 221  
Qy 143 rLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIl 163  
Db 222 AATCATGACTTTTGGAGGAATGATGGTGCCATCTTTCAGCGGAAAGTTGCAGATCTCAT 281  
Qy 163 eGlyArgGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIl 183  
Db 282 GGGTCGAAAAGGGACTATGTGGTTTGTCTCAAATTTTCTGCATCTCTCGGTTGGTTGCAGT 341  
Qy 183 eSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyVa 203  
Db 342 AGCATTAGCAAAAGACTCCATGTGGCTTGATATTGGAAGACTATCCACAGGATTTGCAGT 401  
Qy 203 lGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetAr 223  
Db 402 TGGTTTATTAAAGCTATGTGATACCAGTTTACATTGCAGAAATAACACCAACATGTTTCG 461  
Qy 223 gGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLe 243  
Db 462 AGGAGCGTTTGTATTGTCTAATCAGCTGATGCAGAGTTGTGGATGTCTTTATTCTACGT 521  
Qy 243 uLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMe 263  
Db 522 CATTCGAAATTTGTTTCATTGGCGTAACTTGGCCTTAATCGGTCTCATTCATGTGCGTT 581  
Qy 263 tLeuIleProGlyLeuPheIleProGluSerProArgTrpLeuAlaLysMetAsnLe 283  
Db 582 GCAAGTTGTGACTTTGTTCTTTTATTCAGAGTCCCTTAGACTACTGGGAAATGGGGACA 641  
Qy 283 uThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrTh 303  
Db 642 TGA AAAAGAAATGTAGAGCTTCATTCGAAAGCTTCGCGGAGATGATGCAGATATCTCTGA 701  
Qy 303 rGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPh 323  
Db 702 AGAAGCCAACTACTATCAAAGAAACCATGATCTTTGTTGATGAAGACCAAAATCGCGGT 761  
Qy 323 eGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGlyLeuLeuVa 343  
Db 762 TATGGATTGTTTCAGAGAAGATATGCTCCATCTGTGTTGTTATGTTGGTGGGACTAATGCT 821  
Qy 343 lLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAl 363  
Db 822 TCTACAACAACCTCTCTGGAAGCTCAGGACTTATGTACTATGTGCGTAGCGTATTGTATA 881  
Qy 363 aAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAl 383  
Db 882 AGGAGGGTTTCCAAGCAGC---ATTGGCTCAATGATTCTTTCAGTGATCATGATACCAA 938  
Qy 383 aThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSe 403  
Db 939 AGCTCTATTGGGTCTGATTTTGGTTGAGAAAAATGGGACGAGACCACTTCTATTGGCCTC 998  
Qy 403 rThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAs 423  
Db 999 TACCGGTGGAATGTGCTTTTTCAGCTTGCTCTCCTCAGTTTTTCTCTCTGCTTTCGG---- 1053  
Qy 423 nIleSerGlnAspSerAsnSerTyrTyrIleLeu-----ThrMetIleSerLe 439  
Db 1054 -----TCATATGGCATGCTTGATGAGCTCACTCCGATTTTCACATG 1094  
Qy 439 uValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLe 459  
Db 1095 TATCGGTGTAGTGGGTTTCATCTCTTCATTTGCCGTAGGCATGGGAGGCTTACCATGGAT 1154  
Qy 459 uMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLe 479  
Db 1155 CATCATGTCTGAGATATTCCCAATGAATGTTAAAGTTTCTGCTGGGACTCTGTTACCTT 1214

Qy	479	uAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSe	499
Db	1215	AGCCAACTGGTCCTTTGGTTGGATTGTTGCTTCGCCCTACAACTTCATGCTAGAGTGGAA	1274
Qy	499	rValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValVa	519
Db	1275	CGCATCAGGAACGTTCTTGATCTCTTTACTATATATGGTGGCGGTATAGTCTTTATTATA	1334
Qy	519	lLeuTrpValProGluThrLysGly	527
Db	1335	TGCGATGGTACCCGAAACTAAAGGA	1359

Search completed: April 13, 2005, 15:31:04  
Job time : 668.423 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 13:36:34 ; Search time 223.162 Seconds  
(without alignments)  
4054.730 Million cell updates/sec

Title: US-10-051-909-36  
Perfect score: 2779  
Sequence: 1 PSSSSSFRPAGKKKKKNQG.....TIFVLSIQRLQWLPECLS 553

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+\_p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074038\_14080/app\_query.fasta\_1.1678  
-DB=Issued Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909\_@CGN\_1\_1\_141\_@runat\_13042005\_074038\_14080 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	578	20.8	5228	4	US-09-919-039-216 Sequence 216, App
2	560.5	20.2	5227	4	US-09-919-172-79 Sequence 79, Appl
3	557.5	20.1	2089	3	US-09-291-922-25 Sequence 25, Appl
4	557	20.0	2017	3	US-09-291-922-21 Sequence 21, Appl
5	551.5	19.8	3915	4	US-09-023-655-1104 Sequence 1104, App
6	542.5	19.5	1853	3	US-09-291-922-23 Sequence 23, Appl
7	528	19.0	2592	3	US-09-591-025-8 Sequence 8, Appl
8	527.5	19.0	1914	3	US-09-291-922-19 Sequence 19, Appl
9	524	18.9	2592	4	US-09-894-927B-8 Sequence 8, Appl
10	523.5	18.8	2856	3	US-09-643-597-135 Sequence 135, App
11	523.5	18.8	2856	4	US-09-480-884A-135 Sequence 135, App
12	523.5	18.8	2856	4	US-09-542-615A-135 Sequence 135, App

13	523.5	18.8	2856	4	US-09-606-421B-135	Sequence 135, App
14	523.5	18.8	2856	4	US-09-221-107-135	Sequence 135, App
15	523.5	18.8	2856	4	US-09-466-396A-135	Sequence 135, App
16	523.5	18.8	2856	4	US-09-476-496A-135	Sequence 135, App
17	523.5	18.8	2856	4	US-09-630-940B-135	Sequence 135, App
18	523.5	18.8	2856	4	US-09-285-479-135	Sequence 135, App
19	522	18.8	1872	3	US-09-291-922-27	Sequence 27, Appl
20	519	18.7	3000	2	US-08-928-692-9	Sequence 9, Appl
21	519	18.7	3000	3	US-09-339-972-9	Sequence 9, Appl
22	514.5	18.5	1431	4	US-09-489-039A-4762	Sequence 4762, Ap
23	497	17.9	1545	4	US-09-489-039A-4731	Sequence 4731, Ap
24	470.5	16.9	1752	4	US-09-679-686B-17	Sequence 17, Appl
25	469	16.9	1776	4	US-09-679-686B-11	Sequence 11, Appl
26	455	16.4	1960	4	US-09-679-686B-15	Sequence 15, Appl
27	451	16.2	1675	4	US-09-679-686B-1	Sequence 1, Appl
28	450	16.2	1704	4	US-09-614-221A-96	Sequence 96, Appl
29	447.5	16.1	1668	4	US-09-614-221A-420	Sequence 420, App
30	446.5	16.1	1704	4	US-09-614-221A-534	Sequence 534, App
31	446	16.0	1506	4	US-09-489-039A-4560	Sequence 4560, Ap
32	442	15.9	1659	4	US-09-248-796A-6610	Sequence 6610, Ap
33	439.5	15.8	987	4	US-09-710-279-1941	Sequence 1941, Ap
34	439.5	15.8	3065	4	US-09-710-279-3697	Sequence 3697, Ap
35	439.5	15.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl
36	435	15.7	740	4	US-09-949-016-3739	Sequence 3739, Ap
37	435	15.7	1566	4	US-09-489-039A-2378	Sequence 2378, Ap
38	421.5	15.2	1626	4	US-09-614-221A-521	Sequence 521, App
39	420	15.1	1515	4	US-09-248-796A-6607	Sequence 6607, Ap
40	418	15.0	1714	4	US-09-774-528-437	Sequence 437, App
41	417	15.0	1943	4	US-09-774-528-168	Sequence 168, App
42	416.5	15.0	1689	4	US-10-162-012-45	Sequence 45, Appl
43	416.5	15.0	4167	4	US-09-949-016-4759	Sequence 4759, Ap
44	416.5	15.0	4385	4	US-10-162-012-43	Sequence 43, Appl
45	413.5	14.9	1521	4	US-09-248-796A-5972	Sequence 5972, Ap

ALIGNMENTS

RESULT 1

US-09-919-039-216  
; Sequence 216, Application US/09919039  
; Patent No. 6727066  
; GENERAL INFORMATION:  
; APPLICANT: Kaber, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
; FILE REFERENCE: PA-0035 US  
; CURRENT APPLICATION NUMBER: US/09/919,039  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/222,113  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 401  
; SOFTWARE: PERL Program  
; SEQ ID NO 216  
; LENGTH: 5228  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6727066 410785.1  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 4929, 4935, 4940, 4945, 4974, 4993  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-919-039-216  
  
Alignment Scores:  
Pred. No.: 2.13e-52 Length: 5228  
Score: 578.00 Matches: 152  
Percent Similarity: 49.22% Conservative: 101  
Best Local Similarity: 29.57% Mismatches: 167  
Query Match: 20.80% Indels: 94  
DB: 4 Gaps: 15  
  
US-10-051-909-36 (1-553) x US-09-919-039-216 (1-5228)







Db 317 ACGTCCGACTGGATCGCGCGCGTTCACCATCGTCTTCGCCCGCGTCATCTTCTTCGGG 376  
QY GlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeu 198  
Db 377 GCGGCCCTCATCATGGCTTCTCCGTCAACTAGCCCATGCTCATGTTCGGGCGCTTCGTG 436  
QY GluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSer 218  
Db 437 GCGGCATCGCGGTGGGTACGCTCTCATATCGCGCCCGTGAACACGGCGAGGTGTCC 496  
QY ProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIle 238  
Db 497 CCGCGTCTGCCCGTGGGTTCACATCCTTCCCGGAGGTGTTCATCAACTTCGGGCATC 556  
QY PheLeuAlaTyrLeu-----LeuGlyMetPheIleProTyrArg 251  
Db 557 CTCCTCGGATATGCTCCAACTTCGCCCTTCGCCCGCCTCTCCCTCCGCTCGGCTGGCGC 616  
QY LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271  
Db 617 ATTATGCTCGGCATAGCGCGGTGCGCTCGCTCGCTCGCTCGCTCGCTCGGCGATG 676  
QY ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291  
Db 677 CCGGAGTCTCCCGCTGGCTCGTCATGAAGGCCCGTCTCGCGGACGCCAAGGTGTGCTT 736  
QY GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311  
Db 737 GCCAAGACGTCCGACACGCCGGAAGAGGCCGCGAGCGCATCGCGACATTAAAGACTGCC 796  
QY 312 Val-----Ala 313  
Db 797 GCGGCATCCCTCTGGGCCTCGACGGCGAGTGGTCCCCGTGCCCAAAACAAAGGAAGC 856  
QY SerSerSerLysArgThr-----ThrIleSerPheGlnGlu 325  
Db 857 AGCGAGGAGAGCGCGTTTGAAGGACCTCATCTGTACCGCATAGCCATGCGCCAC 916  
QY LeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGln 345  
Db 917 ATC-----CTCATCGGGGAATCGGCATCCACTTCTTCCAG 952  
QY AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365  
Db 953 CAGTCTTCGGGCATCGACGCGTCTGTCTACAGCCCGTACTGTTTCAAGAGCGCGGC 1012  
QY ValThrAsnSerAspLeu-----AlaThrCysSerLeuGlyAlaIleGlnValLeu 382  
Db 1013 ATCAGGGCGACAGCCGTCTCGCGGCACACCGTGGCGGTGGGGCCACCAATACGGTC 1072  
QY AlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIle 402  
Db 1073 TTCATCCTGGTGGCCACCTTCTCCTCGACCGCATCCGCCGCGCGCGCTGGTGTGACC 1132  
QY SerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAsp 422  
Db 1133 AGCAGCGCGGCATGCTCGTCTCCTTAGTGGGCTCGCGACGGGGCTCACCGTCAATCAGC 1192  
QY AsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIle 442  
Db 1193 CGCCACCCGACGAGAAGATCACCTGGGGCATC-----GTCCTGTGCATCTTCTGCATC 1246  
QY ValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTyrTrpLeuMetMetSer 462  
Db 1247 ATGGCCTACGTGGCCTTCTTCTCATCGGCTCGGCCCATCATCATGGGTGTACAGTCTG 1306  
QY GluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrp 482  
Db 1307 GAGATCTTCCCGCTGCACGTGCGCGCTGGGCTGCTCCCTGGGCGTGGCCGTCACACCGC 1366  
QY LeuThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLeuThrTrpSerValGly 501

Db 1367 CTGACCAGCGCGTGTATCTCCATGACCTTCAATTCGCTGTCCAAGGCCATGACCATCGGC 1426  
QY GlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrp 521  
Db 1427 GGCGCCTTCTTCTTTCGCGGCATCGCCTCATTCGCATGGGTGTCTTCTTTCGCCTAC 1486  
QY 522 ValProGluThrLysGly 527  
Db 1487 CTGCCGAGAGACCCGCGC 1504  
RESULT 4  
US-09-291-922-21  
; Sequence 21, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; EARLIER FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 21  
; LENGTH: 2017  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-09-291-922-21  
Alignment Scores:  
Pred. No.: 8.31e-51 Length: 2017  
Score: 557.00 Matches: 148  
Percent Similarity: 49.01% Conservative: 99  
Best local Similarity: 29.37% Mismatches: 211  
Query Match: 20.04% Indels: 46  
DB: 3 Gaps: 9  
US-10-051-909-36 (1-553) x US-09-291-922-21 (1-2017)  
QY 80 SerSerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSer--- 98  
Db 75 TCCTCCTCTGCACCAACCGGAGATGGCTTCCGCCCGCTGCCGAGGCCGTCCGCCGAAG 134  
QY 99 -----AlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111  
Db 135 AAGAAGGCAACGTCGCGTTCGCTTCGCCCTGCGCATCCTCGCCTCCATGACCTCCATC 194  
QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131  
Db 195 CTCCTCGGCTACGATATCGGGGTGATGAGCGGGCGTCTGTGTACATCAAGAAGGACTTC 254  
QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151  
Db 255 AACATCAGTGACGGGAAGGTGGAGGTTCTCATGGGCATACTGAACCTTACTCGCTCATC 314  
QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171  
Db 315 GGCTCCTTCGCGGGCGGACGTCGGACTGGATCGGCGCGGTACACCATCGTGTTC 374  
QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191  
Db 375 GCCGCGCTCATATTCTTCGGGGGSGGTTCCTCATGGGGTTCGCCGTCAACTACGCCATG 434  
QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211  
Db 435 CTATGTTCGGCCGCTTCGTGGGCGGCATCGGCGTGGGTACGCGCTCATGATCGGCGCG 494  
QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231



Db 495 GTGTACACCGCGGAGGTGTGCGCGCGTGGCTTCCTGACGTCGTTCCCGGAG 554

Qy 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeu-----Leu 244

Db 555 GTGTTCACTAACTTCGGCATCCTGCTCGGTACGTCTCGAACTATGCTTTCTCCCGCTTG 614

Qy 245 GlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeu 264

Db 615 CCGCTGAACCTCGGTGGCGCATCATGCTCGGCATCGGCGCGCGTCCGTGCTGCTC 674

Qy 265 IleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThr 284

Db 675 GCGCTCATGCTGCTCGGCATGCCGGAGTCGCCCGGTGGCTGGTATGAAGGACGCCCT 734

Qy 285 GluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGlu 304

Db 735 GCGGACGCCAAGGTGGTGTGGAGAGACCTCCGACACGCGGAGGAGCGCGGAGCGC 794

Qy 305 ValAsnAspIleLysArgAla-----ValAlaSer 314

Db 795 CTGGCCGACATCAAGCCGCGCGCATCCCTGAGGAGCTCGACGCGGACGTGGTGACC 854

Qy 315 SerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrPro 334

Db 855 GTCCCAAGAGAGGAGCGGAAACGAGAGCGGGTGTGGAAGGAGCTCATCTGTGCCCG 914

Qy 335 -----LeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeu 347

Db 915 ACCCGGCCATCGCGGCATCCTGTCTCGGGATCGGCATCCACTTCTCCAGCATGCG 974

Qy 348 SerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThr 367

Db 1035 AACGACAAACACTTCTTGGGCACCACCTTGGCCGTCGGTGTCAACCAAGAGCTTTTCATC 1094

Qy 385 GlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThr 404

Db 1095 TTGTTGGCACTTCTTCATCGACGGCTCGGCGGGCGCGCTGTGCTGGGCAGCAG 1154

Qy 405 SerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsnIle 424

Db 1155 GCGGGATAATCTCTCCCTCATCGGCCTCGGCGCGGCTCACCGTC-----GTC 1205

Qy 425 SerGlnAspSerAsnSerTyrTyrIleLeuThrMet---IleSerLeuValGlyIleVal 443

Db 1206 GGCAGCACCCGACGCCAAGATACCTTGGGCCATCGGCCTAAGCATCGCCTCCACCCTC 1265

Qy 444 SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGlu 463

Db 1266 GCCTACGTGCTTCTTCCATCGGCCCTGGCCCATCATCACGTGGGTGTACAGCTCGGAG 1325

Qy 464 IleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeu 483

Db 1326 ATCTTCCCGCTCCAGGTGCGCGCTGGGTGCTCGCTCGGCGTCCGCCCAACCGCGTC 1385

Qy 484 ThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLeuThrTrpSerValGlyGly 502

Db 1386 ACCAGCGGCGTCATCTCCATGACCTTCTGTCGCTGTCCAAGGCCATCACCATCGCGCGC 1445

Qy 503 ThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpVal 522

Db 1446 AGCTTCTCTCTACTCCGGCATCGCGCGCTCGCCTGGGTGTTCTTCTACACCTACCTC 1505

Qy 523 ProGluThrLysGly\*\*AsnSerArg-----GlyAspThrIlePhe 536

Db 1506 CCGGAGACCCGCGCGGACCGCTGGAGGAGATGAGCAAGCTGTTTCGGCGACACGCGCGC 1565

Qy 537 valSerLeuSer 540

Db 1566 GCCTCGGAATCA 1577

RESULT 5

US-09-023-655-1104

; Sequence 1104, Application US/09023655

; Patent No. 6607879

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

; TITLE OF INVENTION: EXPRESSION

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/023,655

; FILING DATE: HERewith

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Zeller, Karen J.

; REGISTRATION NUMBER: 37,071

; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO: 1104:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3915 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GENBANK

; CLONE: g183684

US-09-023-655-1104

Alignment Scores:

Pred. No.: 1.08e-49 Length: 3915

Score: 551.50 Matches: 145

Percent Similarity: 48.95% Conservative: 87

Best Local Similarity: 30.59% Mismatches: 180

Query Match: 19.85% Indels: 62

DB: 4 Gaps: 13

US-10-051-909-36 (1-553) x US-09-023-655-1104 (1-3915)

Qy 94 GluSerHisValSerAlaPheLeuCys-ThrLeuIleValAlaLeuGlyProIleGlnPh 113

Db 254 GAAGGTCACCCAGCTCTGATATTGCCATCACAGTTGTCATCGGCTCTTCCAATT 313

Qy 113 eGlyPheThrSerGlyPhe----- 119

Db 314 TGGCTACAACACTGGGGTCATCAATGCTCTGAGAAGATCATAAAGGAATTATCAATAA 373

Qy 120 -----SerSerProThrGlnAspAlaMetValArgAspLeu---As 132

Db 374 AACTTTGACGGACAAGGAAATGCCCCACCCTCTGAGGTGCTGCTCAGCTCTCTGGTC 433

Qy 132 nLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGl 152

Db 434 CTTGTCTGTGGCCATATTTTCC-----GTGGGGGTATGATCGG 472  
QY 152 YAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAl 172  
Db 473 CTCCTTTTCGGTCGGACTCTTCGTCAACCGCTTTGGCAGGCGCAATTCAATGCTGATTGT 532  
QY 172 aAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPhe-----AlaLysAspAl 189  
Db 533 CAACCTGTGTGGCTGTCACTGGTGGCTGCTTTATGGGACTGTGTAAGTAGTAAAGTCGGT 592  
QY 189 aSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrTh 209  
Db 593 TGAATGCTGATCCTGGTGCCTTGGTTATTGGCCTCTTCTGCGGACTTGCACAGGTTT 652  
QY 209 rValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVa 229  
Db 653 TGTGCCCATGTACATTGGAGAGATCTGCCTACTGCCCTGCGGGTGCCTTTGGCACTCT 712  
QY 229 lAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMet---PheIl 248  
Db 713 CAACAGCTGGGCATCGTTGTTGGAATTCTGGTGGCCAGATCTTTGGTCTGGAATTCAT 772  
QY 248 e-----ProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMe 263  
Db 773 CCTTGGTCTGAAGAGCTATGGCCGCTGCTACTGGGTTTACCATCCTCTCCTGCTATCCT 832  
QY 263 tLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLe 283  
Db 833 ACAAGTGCAGCCCTTCCATTTTGCCTGAAAGTCCCAGATTTTGTCTC---ATTAACAG 889  
QY 283 uThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPhe-----GluThrAspIl 301  
Db 890 AAAAGAAGAG--GAGAATGCTAAGCAGATCCTCCAGCGTTGTGGGGCACCCAGGATGT 946  
QY 301 eThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrTrIl 321  
Db 947 ATCCAAAGACATCCAGGAGATGAAAGATGAGAGTGCAAGGATGTCAAGAAAAAGCAAGT 1006  
QY 321 eSerPheGlnGluLeu---AsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGl 340  
Db 1007 CACCGTGTAGAGCTCTTTAGAGTGTCCAGCTACCGACAGCCCATCATCTTTCCATTGT 1066  
QY 340 yLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIl 360  
Db 1067 GCTCCAGCTCTCTCAGCAGCTCTCTGGGATCAATGCTGTGTTCTATTACTCAACAGGAAT 1126  
QY 360 ePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGl 380  
Db 1127 CTTCAGGATGCAGGTGTTCAAGAGCCCATCTATGCCACCATCGGCGCGGTGTGTTAA 1186  
QY 380 nValLeuAlaThrGlyValThrThrTriPleuLeuAspArgAlaGlyArgArgIleLeuLe 400  
Db 1187 TACTATCTTCACTGTAGTTTCTCTATTCTGGTGGAAAGGCGCAGGAAGGACTCTGCA 1246  
QY 400 uIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLe 420  
Db 1247 TATGATAGGCCTTGGAGGGATGGCTTTTGTCCACGCTCATGACTGTTCTTTGTTATT 1306  
QY 420 uLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuVa 440  
Db 1307 AAAGGATAAC-----TATAATGGGATGAGCTTTGTCTCTGTTATT 1345  
QY 440 lGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMe 460  
Db 1346 GGCTATCTTGGTCTTGTAGCCTTCTTTGAAATTGGACCAGGCCCATTCCTGGTTTAT 1405  
QY 460 tMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAl 480  
Db 1406 TGTGGCCGAACCTTTCAGCCAGGCGCCCGCCAGCTGCGATGGCAGTGGCCGCGCTGCTC 1465  
QY 480 aAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVa 500

Db 1466 CAACTGGACTCCAACTTCCTAGTCGGA-----TTGCTCTTCCCTC 1507  
QY 500 lGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThr----- 514  
Db 1508 CGCTGCTCACTATTATTAGGAGCCTACGTTTATTATCTTCACCGGCTTCCTCATTACCTT 1567  
QY 515 -IleValPheValValLeuTrpValProGluThrLysGly 527  
Db 1568 CTTGGCTTTTACCTTCTTCAAAGTCCCTGAGACCCCGTGGC 1607  
RESULT 6  
US-09-291-922-23  
; Sequence 23, Application US/09291922  
; Patent No. 6383776  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB-1163  
; CURRENT APPLICATION NUMBER: US/09/291,922  
; CURRENT FILING DATE: 1999-04-14  
; EARLIER APPLICATION NUMBER: 60/083,044  
; EARLIER FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 23  
; LENGTH: 1853  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-291-922-23  
Alignment Scores:  
Pred. No.: 2.86e-49 Length: 1853  
Score: 542.50 Matches: 137  
Percent Similarity: 50.64% Conservative: 102  
Best local Similarity: 29.03% Mismatches: 183  
Query Match: 19.52% Indels: 51  
DB: 3 Gaps: 9  
US-10-051-909-36 (1-553) x US-09-291-922-23 (1-1853)  
QY 93 ArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGln 112  
Db 239 CGCAAAAGGAACAAGTATGCTTTTGTGTGCTATGCTGGCCTCCATGACTTCCATCTTG 298  
QY 113 PheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAsn 132  
Db 299 CTTGGTTATGATATTGGAGTGATGATGGAGAGCCATATACATAAAAGGGACCTGAAA 358  
QY 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152  
Db 359 GTCTCGGACGCAAAATCGAGATCCTGCTCGGAATCATCAACCTATATACTCTCTGATAGGC 418  
QY 153 AlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAla 172  
Db 419 TCATGTCTCGCGGCAGAAACCTCCGACTGGATAGGTCCCGTTACACGATTGTTTTCGCC 478  
QY 173 AlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeu 192  
Db 479 GGCACCATCTTCTTGTCTGGAGCACITTCATCATGGGTTTCTCCCAATATTTCCTTCTC 538  
QY 193 TyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProVal 212  
Db 539 ATGTTTGGCCGTTTCGTGCTGGCATTTGGCATCGGCTACGCCCTCATGATAGCCCCCGTC 598  
QY 213 TyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeu 232  
Db 599 TACACCGCCGAGGTCTCCCGGCTCCTCTCGTGGCTTCTTCTTCTTCTTCCCTGAGGTA 658  
QY 233 SerValThrPheGlyIlePheLeuAlaTyrLeu-----LeuGly 245

Db 659 TTTATTATGAGGGATATTAAATTGGATACATATCAAACTATGCATTTTCGAAGCTGACA 718  
Qy 246 MetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle 265  
Db 719 CTAAGGTGGATGGCAATGATGCTTGAGTTGGTGAATACCTTCGGTACTCCTAACA 778  
Qy 266 ProGlyLeuPheIleProGluSerProArgTrpLeuAlaLys----- 280  
Db 779 GTAGGAGTGTGGCGATGCCGGAGTCCCAAGTGGCTTGTGATGAGGGTTCGTTTGGGA 838  
Qy 281 -----MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295  
Db 839 GAGGCAAGAAAAGTGCTTAACAAAACCTCAGACAGCAAGGAAGAGCCCAACTA----- 892  
Qy 296 GlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaVal-AlaSerSe 315  
Db 893 -----AGGCTAGCGGAANTCAAAACAAGCCGACGGGATCCCC 928  
Qy 315 rSerLysArgThrThrIleSerPheGlnGluLeuAsnGln----- 328  
Db 929 GAGAGTTGCAACGACGACGTCGTTTCAGGTAAATAAACAAAGCAACGGTGAAGGTGTATGG 988  
Qy 329 -----LysLysTyrArgThrPro-----LeuLeuLeuGlyIleGl 340  
Db 989 AAAGAGCTCTTCCTC-TATCCAACGCCCGCAATTTCGTCACATCGTAATCGTGCCTTGG 1047  
Qy 340 yLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIl 360  
Db 1048 TATTCACCTCTTCCAACAAGCGTCGGGGGTAGACGCCGTCGTTTGTACAGCCCCAGGAT 1107  
Qy 360 ePheLysAlaAlaGlyValThrAsnSerAsp-----LeuAlaThrCysSerLeuGl 377  
Db 1108 CTTCGAAAGGCTGGGATTACAAACGACGACGCATAAGCTTCTTGGAACCGTGGCCGTGG 1167  
Qy 377 yAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgAr 397  
Db 1168 ATTCGTTAAGACCGTGTTTCATCTTGGCGGCTACGTTTACGTTGGACCGCGTGGGTCGTCG 1227  
Qy 397 gileLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValva 417  
Db 1228 TCCGTTGTATTGTCTAGTGTCCGGCGCATGGTGCTCTCGCTTCTCACGCTTGGCATCAG 1287  
Qy 417 lPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIl 437  
Db 1288 CCTCACTGTTATTGATCAT---TCGGAGAGGAAATTAATGTGGGCCGTTGGATCGAGCAT 1344  
Qy 437 eSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIlePr 457  
Db 1345 AGCCATGGTG-----TTGGCTTACGTGGCCACGTTCTCCATCGGTGCGGGTCCCATCAC 1398  
Qy 457 oTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAl 477  
Db 1399 GTGGGTCTATAGTTCTGAGATCTTCCCGTTGAGGCTGCGGGCGCARGGTGCGGCCGCGGG 1458  
Qy 477 aThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLe 496  
Db 1459 AGTTGCGGTGAATAGGACCACCTAGCGCGGTTGTCTCAATGACTTTTCTGTCCCTCACTAG 1518  
Qy 496 uThrTrpSerValGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleVa 516  
Db 1519 AGCCATCACTATTGGTGGAGCTTTCTTCCTTTATTGTGGCATTGCTACTGTTGGTGGAT 1578  
Qy 516 lPheValValLeuTrpValProGluThrLysGly 527  
Db 1579 ATTCTTTTACACCGTCTTGCCTGAGACCCGGGGA 1612

RESULT 7

US-09-591-025-8  
; Sequence 8, Application US/09591025  
; Patent No. 6303373  
; GENERAL INFORMATION:  
; APPLICANT: Bogan, Jonathan S.  
; APPLICANT: Lodish, Harvey F.

; TITLE OF INVENTION: Method of Measuring Plasma Membrane  
; TITLE OF INVENTION: Targeting of GLUT4  
; FILE REFERENCE: 0399.1210-004  
; CURRENT APPLICATION NUMBER: US/09/591,025  
; CURRENT FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/154,078  
; PRIOR FILING DATE: 1999-09-15  
; PRIOR APPLICATION NUMBER: 60/138,237  
; PRIOR FILING DATE: 1999-06-09  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 2592  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences  
US-09-591-025-8

Alignment Scores:  
Pred. No.: 2.06e-47 Length: 2592  
Score: 528.00 Matches: 131  
Percent Similarity: 52.58% Conservative: 83  
Best Local Similarity: 32.19% Mismatches: 165  
Query Match: 19.00% Indels: 28  
DB: 3 Gaps: 7

US-10-051-909-36 (1-553) x US-09-591-025-8 (1-2592)

Qy 133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152  
Db 562 CTCTCCGTGGCCATCTTTTCC-----GTGGGCGGCATGATTTCC 600  
Qy 153 AlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAla 172  
Db 601 TCCTTCCTCATTTGGTATCATCTCTCAGTGGCTTGAAGGAAAAGGCCATGCTGGTCAAC 660  
Qy 173 AlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSer----- 190  
Db 661 AATGTCTGGCGGTGTGGGGGCGAGCCCTCATGGGCGCTGGCCAACGCTGCTGCCTCCTAT 720  
Qy 191 ---PheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209  
Db 721 GAAATGCTCATCTTGGACGATTCTCTCATTTGGCGCCTACTCAGGGCTGACATCAGGGCTG 780  
Qy 210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229  
Db 781 GTGCCCATGTACGTGGGGAGATTGCTCCCACTCACCTGCGGGGCGCCCTGGGAGCGCTC 840  
Qy 230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyr-----Leu 243  
Db 841 AACCAACTGGCCATTGTTATCGCATTTCTGATCGCCCGAGGTGTGGGCTTGGAGTCCCTC 900  
Qy 244 LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263  
Db 901 CTGGCACTGCCAGCCTGTGGCCACTGCTCTCTGGGCCCTCACAGTGCTACCTGCCCTCCTG 960  
Qy 264 LeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeu---AlaLysMetAsn 282  
Db 961 CAGCTGGTCTGTGCTGCCCTTCTGTCTCCGAGAGCCCCCGCTACCTCTACATCATCCAGAAT 1020  
Qy 283 LeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu-----ThrAsp 300  
Db 1021 CTCGAGGGGCTGCCAGAAAGAGTCTGAAGCGCCTGACAGGCTGGGCGCGATGTTTCTGGA 1080  
Qy 301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThr 320  
Db 1081 GTGCTGGCTGAGCTGAAGGATGAGAAGCGGAAGCTG-----GAGCGTGAGCGGCCACTG 1134  
Qy 321 IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGly 340  
Db 1135 TCCCTGCTCCAGCTCCTGGGCGAGCCGTACCCACCGGACGCCCTGATCATTTGGGTCGTG 1194



QY	341	LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerile	360
DB	1195	CTGCAGCTGAGCCAGCAGCTCTCTGGCATCAATGCTGTTTCTATTATTCGACCAGCATC	1254
QY	361	PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln	380
DB	1255	TTCGAGACAGCAGGGGTAGCCAGCCTGCCTATGCCACCATTAGGAGCTGGTGGTCAAC	1314
QY	381	ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu	400
DB	1315	ACAGTCTTACCTTGGTCTCGGTGTTGTTGGTGAGCGGGGGCGCGGACGCTCCAT	1374
QY	401	IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPheLeu	420
DB	1375	CTCCTGGGCCCTGGCGGGCATGTGGCTGTGCCATCCTGATGACTGTGGCTCTGCTCCTG	1434
QY	421	LysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuVal	440
DB	1435	CTGAGCGAGTTCCA-----GCCATGAGTACGTCTCCATTGTG	1473
QY	441	GlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMet	460
DB	1474	GCCATCTTTGGCTTCGTGGCATTTTGTAGATTGGCCCTGCCCATTCCTTGGTTCATC	1533
QY	461	MetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAla	480
DB	1534	GTGGCCGAGCTCTTCAGCCAGGACCCCGCCGAGCATGGCTGTGGCTGGTTTCTCC	1593
QY	481	AsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVal	500
DB	1594	AACCTGGACGAGCAACTTCATCATTTGGCATGGGTTTCCAGTATGTTGCGGAGGCTATGGGG	1653
QY	501	GlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeu	520
DB	1654	CCCTACGCTTCTCTTCTATTTTGGGTCCCTCCTGCTGGGCTTCTTCATCTTCACCTTCTTA	1713
QY	521	TrpValproGluThrLysGly	527
DB	1714	AGAGTACCTGAAACTCGAGGC	1734

```

RESULT 8
US-09-291-922-19
; Sequence 19, Application US/09291922
; Patent No. 6383776
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB-1163
; CURRENT APPLICATION NUMBER: US/09/291,922
; CURRENT FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 60/083,044
; EARLIER FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Zea mays
US-09-291-922-19

```

Alignment Scores:		
Pred. No.:	1.37e-47	1914
Score:	527.50	132
Percent Similarity:	48.72%	97
Best Local Similarity:	28.09%	186
Query Match:	18.98%	55
DB:	3	7
	Length:	
	Matches:	
	Conservative:	
	Mismatches:	
	Indels:	
	Gaps:	

US 10-051-909-36 (1-553) x US-09-291-922-19 (1-1914)

[illegible]

```

Db      1160 ACCGTCGGGGGCATCACCCGGACACCAAGGTCGGTGGGCCGTC-----GCCCTGTGC 1213
Qy      439 LeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrp 458
Db      1214 ATCGCGTCAACCGCTGCTTACATCGCCTTCTTCCATCGGCCTCGGSCCATCACGGGC 1273
Qy      459 LeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThr 478
Db      1274 GTGTACACCTCGGAAATATTCCCGCTGCAGTGCAGCGCTGGCTTCCGCGTGGGTGTG 1333
Qy      479 LeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr---ThrAsnLeuMetLeuThr 497
Db      1334 GCGAGCAACCCGCTCACCGCGCTCATCTCCATGACCTTCCCTGTCCCTCTCCAAGGCC 1393
Qy      498 TrpSerValGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPhe 517
Db      1394 ATCACCATCGGCGCAGCTTCTTCCCTACTCCGGCATCGCGCGGTGGCTTGGTTTTC 1453
Qy      518 ValValLeuTrpValProGluThrLysGly 527
Db      1454 TTCTTACGTGCGCTCCCGGAGACACGCGGC 1483

```

RESULT 9

```

US-09-894-927B-8
; Sequence 8, Application US/09894927B
; Patent No. 6632924
; GENERAL INFORMATION:
; APPLICANT: Bogan, Jonathan S.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: Method of Measuring Plasma Membrane
; TITLE OF INVENTION: Targeting of GLUT4
; FILE REFERENCE: 0399.1210-005
; CURRENT APPLICATION NUMBER: US/09/894,927B
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/591,025
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: US 60/154,078
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: US 60/138,237
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2592
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified GLUT4 containing myc tag sequences
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (2592)
US-09-894-927B-8

```

Alignment Scores:

Pred. No.:	5.7e-47	Length:	2592
Score:	524.00	Matches:	131
Percent Similarity:	52.33%	Conservative:	82
Best local Similarity:	32.19%	Mismatches:	166
Query Match:	18.86%	Indels:	28
DB:	4	Gaps:	7

US-10-051-909-36 (1-553) x US-09-894-927B-8 (1-2592)

```

Qy      133 LeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGly 152
Db      562 CTCTCCGTGGCCATCTTTTCC-----GTGGCGGCATGATTTC 600
Qy      153 AlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAla 172
Db      601 TCCTTCCTATTGGTATCATCTCTCAGTGGCTTGGAAAGGAAAGGCCCATGCTGGTCAAC 660
Qy      173 AlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSer----- 190

```

```

Db      661 AATGCTCTGGCGGTGCTGGGGGGCAGCCTCATGGGCCCTGGCCAACGCTGCTGCCTCCTAT 720
Qy      191 ---PheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThr 209
Db      721 GAAATGCTCATCCTTGGACGATTCTCTATTGGCGCCTACTCAGGCGTGACATCAGGGCTG 780
Qy      210 ValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVal 229
Db      781 GTGCCCATGTAGTGGGGAGATTGCTCCCACTCACCTGCGGGCGCCCTGGGACGCTC 840
Qy      230 AsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyr-----Leu 243
Db      841 AACCAACTGGCCATTGTTATCGGCATTCTGATCGCCCAAGGTGCTGGGCTTGGAGTCCCTC 900
Qy      244 LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet 263
Db      901 CTGGGCACCTGCCAGCCTGTGGCCACTGCTCCTGGGCCCTCACAGTGTACCTGCCCTCCTG 960
Qy      264 LeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeu---AlaLysMetAsn 282
Db      961 CAGTGGTCTGTGCCCTTCTGTCCCGAGAGCCCCCGCTACCTCTACATCATCCAGAAT 1020
Qy      283 LeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGlu-----ThrAsp 300
Db      1021 CTCGAGGGGCTGCCAGAAAGAGTCTGAAGCGCCTGACAGGCTGGGCGCATGTTTCTGGA 1080
Qy      301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThr 320
Db      1081 GTGCTGGCTGAGCTGAAGGATGAGAAGCGGAAGCTG-----GAGCGTGAGCGGCCACTG 1134
Qy      321 IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGly 340
Db      1135 TCCCTGCTCCAGCTCCTGGGCAGCGCTACCCACGGCAGCCCTGATCATTTGGGTCGTG 1194
Qy      341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIle 360
Db      1195 CTGCAGCTGAGCCAGCAGCTCTCTGGCATCAATGCTGTTTTCTATTATTGACCAGCATC 1254
Qy      361 PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln 380
Db      1255 TTCGAGACAGCAGGGGTAGGCCAGCCTGCCTATGCCACCACATAGGAGTGGTGTGTTCAAC 1314
Qy      381 ValLeuAlaThrGlyValThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu 400
Db      1315 ACAGTCTTCACCTTGGTCTCGGTGTGTTGGTGAGCGGGCGGGCGCGGACGCTCCAT 1374
Qy      401 IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeu 420
Db      1375 CTCCTGGGCCCTGGCGGGCATGTGTGGCTGTGCCATCCTGTGACTGTGGCTCTGCTCCTG 1434
Qy      421 LysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuVal 440
Db      1435 CTGGAGCGAGTTCCA-----GCCATGAGCTACGTCTCCATTGTG 1473
Qy      441 GlyIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMet 460
Db      1474 GCCATCTTTGGCTTCGTGGCATTTTGTGAGATTGGCCCTGGCCCCCATTCCTTGGTTTCATC 1533
Qy      461 MetSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAla 480
Db      1534 GTGGCCGAGCTCTTCAGCCAGGAGACCCCGCCCGCAGCCATGGCTGTGGCTGTTTCTCC 1593
Qy      481 AsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerVal 500
Db      1594 AACTGGACGAGCAACTTCATCATTTGGCATGGGTTTCCAGTATGTTGGGAGGCTATGGGG 1653
Qy      501 GlyGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeu 520
Db      1654 CCCTAGCTCTTCTTCTATTGGGGTCTCTCTGTGGGCTTCTTCTCATCTTACCTTCTTA 1713
Qy      521 TrpValProGluThrLysGly 527

```

Db 1714 AGAGTACCTGAAACTCGAGGC 1734

RESULT 10

US-09-643-597-135

; Sequence 135, Application US/09643597

; Patent No. 6426072

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C11

; CURRENT APPLICATION NUMBER: US/09/643,597

; CURRENT FILING DATE: 2000-08-21

; NUMBER OF SEQ ID NOS: 369

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 135

; LENGTH: 2856

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-643-597-135

Alignment Scores:

Pred. No.:	7,68e-47	Length:	2856
Score:	523.50	Matches:	151
Percent Similarity:	45.91%	Conservative:	96
Best Local Similarity:	28.07%	Mismatches:	229
Query Match:	18.84%	Indels:	62
DB:	3	Gaps:	12

US-10-051-909-36 (1-553) x US-09-643-597-135 (1-2856)

QY 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySer-AsnArgG1 46

Db 14 CCGAGTGAGCACGCCAGGAGCAGGAGACCAAACGACGGGGTCCGAGTCAGAGTCGGAG 73

QY 46 yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63

Db 74 TGGAGTCCCGGACCGGAGCACGAGCCTGACGGGAGAGCGCGCTCGCAGCCCGCTCG 133

QY 63 gArgProLeuLeuAsnThrGlySerTriPtyrArgMetSerSerArgGlnSerSerPheAl 83

Db 134 CCACCCCGGTACCCGG-----CGCAGCCAGAGCCACCAGCGCGAGCGCTGC 178

QY 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh 103

Db 179 CATGGAGCCAGCAGCAAGAGCTGACGGGTTCG-----CTCATGTGGTGT 226

QY 103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh 123

Db 227 GGGAGGAGCAGTGCTTGGCTCCCTGCAGTTGGCTACAACACTGGAGTCATCAATGCCCC 286

QY 123 rGlnAspAlaMetValArgaspLeuAsn----- 132

Db 287 CCAGAAGGTGATCGAGAGGTTCTACAACCAGACATGGGTCCACCGCTATGGGAGAGCAT 346

QY 133 -----LeuSerIleSerGluPheSerAlaPheG1 142

Db 347 CCTGCCACCACCGCTCACCACGCTCTGGTCCCTCTCAGTGGCCATCTTTTCT----- 398

QY 142 ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162

Db 399 -----GTTGGGGGCATGATTGGCTCCTTCTCTGTGGGCTTTTCGTTAACCG 445

QY 162 rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182

Db 446 CTTTGGCCGGCGGAATTCAATGCTGATGATGAACCTGTGGCCTTCGTGTCCGCGTGCT 505

QY 182 aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuG1 199

Db 506 CATGGGCTTCTCGAAACTGGSCAAGTCCTTTGAGATGCTGATCCTGGGCGCTTCATCAT 565

QY 199 uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr 219

Db 566 CGGTGTGTACTGCGGCTGACACAGGCTTCGTGCCCATGTATGTGGTGAAGTGTCAACC 625

QY 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239

Db 626 CACAGCCTTTCGTGGGGCCCTGGGCACCTGCACCAGCTGGGCATCGTCGTCGGCATCCT 685

QY 239 eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTrpArgLeuLe 253

Db 686 CATGCCCCAGGTGTTTCGGCCTGGACTCCATCATGGGCAACAAGGACCTGTGGCCCTGCT 745

QY 253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1 273

Db 746 GCTGAGCATCATTTTCATCCCGCCCTGCTGCAGTGCATCGTGTGCCCTTCTGCCCCGA 805

QY 273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291

Db 806 GAGTCCCGCTTCTCTGCTC---ATCAACCGCAACGAGGAGAACCGGGCCAAAGAGTGTGCT 862

QY 291 uGlnValIleuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl 311

Db 863 AAAGAAAGCTGCGCGG---ACAGCTGACGTGACCCATGACCTGCAGGAGATGAAGGAAGA 919

QY 311 aValAlaSerSerLysArgThrThrIleSerPheGlnGluLeu---AsnGlnLysLy 330

Db 920 GAGTCGCAGATGATGCGGGGAGAAAGGTCAACCATCTCTGGAGCTGTTCGCGTCCCCCGC 979

QY 330 sTyrArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyI1 350

Db 980 CTACCCGACGCCCATCTCTCATCGCTGTGGTGTGCTGCAGTGTCCACGAGCTGTCTGGCAT 1039

QY 350 eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs 370

Db 1040 CAACGCTGTCTTCTATTACTCCAGGACATCTTCGAGAAGCGGGGTGCAGCAGCCTGT 1099

QY 370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe 390

Db 1100 GTATGCCACCATTTGGCTCCGGTATCGTCAACACGCGCCTTCACTGTCTGCTGTGTGT 1159

QY 390 uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy 410

Db 1160 GGTGAGGAGCAGCGCCGCGGACCTGCACCTCATAGGCTCGCTGGCATGGCGGGTTG 1219

QY 410 sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430

Db 1220 TGCCATACTCATGACCATCGCGCTAGCACTGTGGAGCAGTACCTGG----- 1268

QY 430 rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450

Db 1269 -----ATGTCCTATCTGAGCATCGTGGCCATCTTTGGCTTTGTGGCCTTCTTGA 1318

QY 450 rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy 470

Db 1319 AGTGGGTCTGCGCCCATCCCATGGTTTCATCGTGGCTGAACCTTTCAGCCAGGGTCCACG 1378

QY 470 sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490

Db 1379 TCCAGCTGCCATTGCCGTTGCAGGCTTCTCCAACCTGGACCTCAAATTCATTGTGGGCAT 1438

QY 490 tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValVa 510

Db 1439 GTGCTTCCAGTATGTGGAGCAACTGTGTGGTCCCTACGCTTTCATCATCTTCACTGTGCT 1498

QY 510 lSerAlaPheThrIleValPheValLeuTrpValProGluThrLysGly 527



\_\_\_\_\_

	INVENTION	TITLE OF INVENTION
1.	COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER	TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
2.	AND DIAGNOSIS OF LUNG CANCER	TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

Length:	2856
Matches:	151
Conservative:	96
Mismatches:	229
Indels:	62
Gaps:	12

US-10-051-909-36 (1-553) x US-09-480-884A-135 (1-2856)

QY	27	ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySer-AsnArgG1	46
DB	14	CCGAGTGAGCAGCCAGGGAGCAGGAGACCAAAACGACGGGGTCCGAGTCAGAGTCGCAG	73
QY	46	YGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr	63
DB	74	TGGGAGTCCCGACCGGAGCAGCAGAGCTGAGCGGGAGAGCGCCGCTCGCACGCCCGCTCG	133
QY	63	gArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl	83
DB	134	CCACCCGGGTACCCG-----CGCAGCCAGAGCCACCAGCGCAGCGGTGC	178
QY	83	aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh	103
DB	179	CATGGAGCCCGCAGCAGCAAGCTGACGGGTGCG-----CTCATGTCTGGGTGT	226
QY	103	rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh	123
DB	227	GGGAGGAGCAGTGCTTGGCTCCCTGCAGTTGGCTACAACACTGGAGTCATCAATGGCCCC	286
QY	123	rGlnAspAlaMetValArgAspLeuAsn-----	132
DB	287	CCAGAAGGTGATCGAGGAGTTCTACAACCAGACATGGGTCCACCGCTATGGGGAGAGCAT	346
QY	133	-----LeuSerIleSerGluPheSerAlaPheG1	142
DB	347	CCTGCCACACCGCTCACCACGCTCTGGTCCCTCTCAGTGGCCATCTTTTCT-----	398
QY	142	ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy	162
DB	399	-----GTTGGGGGCATGATTGGCTCCTTCTCTGTGGGCCCTTTTCGTTAACCG	445
QY	162	rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl	182
DB	446	CTTTGGCCCGCGGAATTCAATGTGCTGATGTAACCTGTGTGGCCTTCGTGTCCGCCGTGCT	505
QY	182	aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuG1	199

US-09-542-615A-135

US-09-542-615A-135

; Sequence 135, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 135  
; LENGTH: 2856  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-542-615A-135

Alignment Scores:  
Pred. No.: 7,68e-47 Length: 2856  
Score: 523.50 Matches: 151  
Percent Similarity: 45.91% Conservative: 96  
Best Local Similarity: 28.07% Mismatches: 229  
Query Match: 18.84% Indels: 62  
DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-542-615A-135 (1-2856)

QY 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySer-AsnArgG1 46  
Db 14 CCGAGTGACGACGCCAGGAGGAGGAGACCAACCGACGGGGTCCGAGTCAGAGTCGCAG 73  
QY 46 YGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63  
Db 74 TGGGAGTCCCGGACCGGAGCAGGAGCCTGAGCGGGAGAGCGCCGCTCGCACGCCGCTCG 133  
QY 63 gArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl 83  
Db 134 CCACCCGGGTACCCGG-----CGCAGCCAGAGCCACCGAGCGCGGTGC 178  
QY 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh 103  
Db 179 CATGGAGCCCGAGCAGCAAGAGTGCAGGGTGC-----CTCATGCTGGCTGT 226  
QY 103 rLeuileValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh 123  
Db 227 GGGAGGAGCAGTGCTTGGCTCCCTGCGAGTTGGCTACAACACTGGAGTCATCAATGCCCC 286  
QY 123 rGlnAspAlaMetValArgAspLeuAsn-----132  
Db 287 CCAGAAGGTGATCGAGGAGTTCTACAACCAAGATGGGTCCACCGCTATGGGGAGAGCAT 346  
QY 133 -----LeuSerIleSerGluPheSerAlaPheG1 142  
Db 347 CCTGCCACCACCGCTCACCACGCTCTGTGTCCTCTCAGTGGCCATCTTTCT----- 398  
QY 142 ySerLeuSerAsnValGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162  
Db 399 -----GTGGGGGCGATGATGGCTCTCTCTGTGGGCTTTTCGTAAACCG 445  
QY 162 rIleGlyArgLysGlySerLeuMetIleAlaIleProAsnIleIleGlyTrpLeuAl 182  
Db 446 CTTTGGCCGGCGGAATCAATGCTGATGATGAACCTGTGGCTTCTGTGGGCTTTTCGTAAACCG 505  
QY 182 aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuG1 199  
Db 506 CATGGGCTTCTCGAAACTGGGCAAGTCCTTTGAGATGCTGATCCTGGGCGCTTCATCAT 565  
QY 199 uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr 219

Db 566 CGGTGTGTACTCGCGCCTGACACAGGCTTCGTGCCCATGTATGTGGTGAAGTGTCAACC 625  
QY 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239  
Db 626 CACAGCCTTTTCGTGGGGCCTTGGGCACCTTCACACAGTGGGCATCGTCGTCGCATCCT 685  
QY 239 eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTrpArgLeuLe 253  
Db 686 CATCGCCAGGTTCGGCCTGGACTCCATCATGGGCAACAAGGACCTGTGGCCCTGTCT 745  
QY 253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProG1 273  
Db 746 GCTGAGCATCATCTTCATCCCGCCCTGCTGCGATGCTGCTGCTGCCCTTCTGCCCGA 805  
QY 273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291  
Db 806 GAGTCCCGCTTCCTGCTC---ATCAACCGCAACGAGGAGAACCGGGCCAAAGAGTGTGCT 862  
QY 291 uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl 311  
Db 863 AAAGAAAGTGCAGCGG---ACAGCTGACGTGACCATGACCTGCAGGAGATGAAGGAAGA 919  
QY 311 aValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeu---AsnGlnLysLy 330  
Db 920 GAGTCGCGAGATGATGCGGAGAGAAGGTTCACCATCTCGGAGCTGTTCGCTCCCGC 979  
QY 330 sTyrArgThrProLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyI1 350  
Db 980 CTACCGCCAGCCCATCTTCATCGCTGCTGTGCTGCTGCTGCCAGCAGCTGTCTGCCAT 1039  
QY 350 eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs 370  
Db 1040 CAACGCTGTCTTCTATTACTCCACGAGCATCTTCGAGAAGCGGGGTGCAGCAGCCTGT 1099  
QY 370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLe 390  
Db 1100 GTATGCCACCATTCGCTCCGCTATCGTCAACACCGCTTCATCTGCTGCTGCTGTTGT 1159  
QY 390 uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy 410  
Db 1160 GGTGGAGGAGCAGGCGCGGACCCCTGCACCTCATAGGCTGCTGCTGCTGCGGTGTG 1219  
QY 410 sLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430  
Db 1220 TGCCATATCTCATGACCATCGCGTACGACTGCTGGAGCAGCTACCTGG----- 1268  
QY 430 rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450  
Db 1269 -----ATGTCCTATCTGAGCATCGTGGCCATCTTTGGCTTTTGGCCTTCTTTGA 1318  
QY 450 rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy 470  
Db 1319 AGTGGTTCCTGGCCCATCCCATGGTTTCATCGTGGCTGAACTCTTCAGCCAGGGTCCACG 1378  
QY 470 sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490  
Db 1379 TCCAGCTGCCATTGCCGTTGAGGCTTCTCCAACTGGACCTCAAATTTTCATGTGGGCAT 1438  
QY 490 tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValVa 510  
Db 1439 GTGCTTCCAGTATGGAGCAACTGTGTGGTCCCTACGCTTTCATCATCTTCACGTGTGCT 1498  
QY 510 lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly 527  
Db 1499 CTTGTTCTGTTCTTTCATCTTCACCTACTTCAAAGTTCCTGAGACTAAAGGC 1550

RESULT 13  
US-09-606-421B-135  
; Sequence 135, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-135

Alignment Scores:
Pred. No.: 7.68e-47 Length: 2856
Score: 523.50 Matches: 151
Percent Similarity: 45.91% Conservative: 96
Best Local Similarity: 28.07% Mismatches: 229
Query Match: 18.84% Indels: 62
DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-606-421B-135 (1-2856)

QY	27	ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySer-AsnArgGl	46
Db	14	CCGAGTGAGCACGCCAGGAGCAGGAGACCAACGACGCGGGTCGGAGTCAGATCGCAG	73
QY	46	yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr	63
Db	74	TGGAGTCCCGGACCGGAGCAGGAGCCTGAGCGGGAGAGCGCCGCTCGCACGCCGCTCG	133
QY	63	qArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl	83
Db	134	CCACCCGCGTACCCGG-----CGCAGCCAGAGCCACCAGCGCAGCGCTGC	178
QY	83	aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh	103
Db	179	CATGGAGCCCGACGACGACGAGCTGACGGGTGCG-----CTCATGTGGTGT	226
QY	103	rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh	123
Db	227	GGGAGGAGCAGTGCTTGGCTCCCTGCGAGTTTGGCTACACACTGGAGTCATCAATGCCCC	286
QY	123	rGlnAspAlaMetValArgAspLeuAsn-----	132
Db	287	CCAGAAGGTGATCGAGGAGTTCTACAAACCAGACATGGGTCCACCGCTATGGGAGAGCAT	346
QY	133	-----LeuSerIleSerGluPheSerAlaPheGl	142
Db	347	CCTGCCCCACACGCTCACCGCTGTGGTCCCTCTCAGTGGCCATCTTTCT-----	398
QY	142	ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy	162
Db	399	-----GTTGGGGGATGATGGCTCCTCTCTGGGCCCTTTTCGTTAACCG	445
QY	162	rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl	182
Db	446	CTTTGGCCGGCGGAATTCATGCTGATGATGAACCTGTGGCCTTCGTGCCCGCGTGT	505
QY	182	aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGl	199
Db	506	CATGGGCTTCTCGAAACTGGGCAAGTCCTTTGAGATGCTGATCCTGGGCCGCTTCATCAT	565
QY	199	uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr	219

Db	566	CGGTGTGTACTGCGGCCTGACCACAGGCTTCGTGCCCATGTATGTGGGTGAAGTGTCAAC	625
QY	219	oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh	239
Db	626	CACAGCCTTTTCGTGGGGCCTTGGGCACCTGCACCAGCTGGGCATCGTCTGGCATCCT	685
QY	239	eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTrpArgLeuLe	253
Db	686	CATGCCCCAGGTGTTCCGGCCTGGACTCCATCATGGGCAACAGGACCTGTGGCCCTGCT	745
QY	253	uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGl	273
Db	746	GCTGAGCATCATCTTCATCCCGCCCTGCTGCAGTGATGATGCTGCTGCCCTTCTGCCCCGA	805
QY	273	uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe	291
Db	806	GAGTCCCGCCTTCCTGCTC---ATCAACCGCAACGAGGAGAACCGGCCAAGAGTGTGCT	862
QY	291	uGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAl	311
Db	863	AAAGAAGCTGCGCGG---ACAGCTGACGTGACCCATGACCTGCAGGAGATGAAGGAAGA	919
QY	311	aValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeu--AsnGlnLysLy	330
Db	920	GAGTCGGCAGATGATCGGGGAGAAGAAGGTCACCATCCTCGAGCTGTTCCGCTCCCCCGC	979
QY	330	sTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIl	350
Db	980	CTACCGCCAGCCCATCCTCATCGCTGTGGTGTGCTGCAGTGCTCCAGCAGCTGTCTGGCAT	1039
QY	350	eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs	370
Db	1040	CAACGCTGTCTTCTATTACTCCACGAGCATCTTCGAGAAGCGGGGTGCAGCAGCCTGT	1099
QY	370	pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe	390
Db	1100	GTATGCCACCATTTGGCTCCGCTATCGTCAACACGCGCCTTCACTGCTGCTGCTGTTGT	1159
QY	390	uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy	410
Db	1160	GGTGGAGCGAGCAGGCGCGGGACCCCTGCACCTCATAGGCCCTCGCTGGCATGGCGGGTTG	1219
QY	410	sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe	430
Db	1220	TGCCATACTCATGACCATCGCGCTAGCACTGTCTGGAGCAGTACCCTGG-----	1268
QY	430	rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe	450
Db	1269	-----ATGTCTCTATCTGAGCATCGTGGCCATCTTTGGCTTTGTGGCCTTCTTTGA	1318
QY	450	rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy	470
Db	1319	AGTGGGTCTTGGCCCCCATCCCATGGTTTCATCGTGGCTGAACCTTTCAGCCAGGGTCCACG	1378
QY	470	sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe	490
Db	1379	TCCAGCTGCCATTGCCGTTGCAGGCTTCTCCAACCTGGACCTCAAATTTCATTGTGGGCAT	1438
QY	490	tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValva	510
Db	1439	GTGCTTCCAGTATGTGGAGCAACTGTGTGGTCCCTACGTTCTTCATCATCTTCACTGTGCT	1498
QY	510	lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly	527
Db	1499	CCTGGTTCTGTTCTTCATCTTCACCTACTTCAAAGTTCTCTGAGACTAAAGGC	1550

RESULT 14
US-09-221-107-135
; Sequence 135, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER



```

; FILE REFERENCE: 210121.455C2
; CURRENT APPLICATION NUMBER: US/09/221,107
; CURRENT FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Human
US-09-221-107-135

Alignment Scores:
Pred. No.: 7.68e-47 Length: 2856
Score: 523.50 Matches: 151
Percent Similarity: 45.91% Conservative: 96
Best Local Similarity: 28.07% Mismatches: 229
Query Match: 18.84% Indels: 62
DB: 4 Gaps: 12

US-10-051-909-36 (1-553) x US-09-221-107-135 (1-2856)

QY 27 ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySer-AsnArgGl 46
Db 14 CCGAGTGCAGCACGCCAGGAGCAGGAGACCAACGACGGGGTCCGAGTCAGAGTCGCAG 73

QY 46 yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr 63
Db 74 TGGGAGTCCCGGACCGGAGCACGAGCCTGAGCGGGAGAGCGCCGCTCGCACGCCCGTCG 133

QY 63 gArgProLeuLeuAsnThrGlySerTrpTyArgMetSerSerArgGlnSerSerPheAl 83
Db 134 CCACCCGGGTACCCGG-----CGCAGCCAGAGCCACCAGCGCAGCGCTGC 178

QY 83 aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh 103
Db 179 CATGGAGCCCGAGCAAGAAAGTTCACACCAGACATGGGTCCACCGCTATGGGGAGAGCAT 226

QY 103 rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerProTh 123
Db 227 GGGAGGAGCAGTGTCTGGCTCCCTGCAGTTGGCTACAACACTGGAGTCATCAATGCCCC 286

QY 123 rGlnAspAlaMetValArgAspLeuAsn----- 132
Db 287 CCAGAAAGGTGATCGAGGAGTTCTACAAACCAGACATGGGTCCACCGCTATGGGGAGAGCAT 346

QY 133 -----LeuSerIleSerGluPheSerAlaPheG1 142
Db 347 CQTGCCACCACCGCTCACCGCTCTGTGCTCCCTCTCAGTGGCCATCTTTTCT----- 398

QY 142 ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162
Db 399 -----GTGGGGGCATGATTGGCTCCTTCTCTGTGGGCCCTTTTCGTAAACCG 445

QY 162 rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182
Db 446 CTTTGGCCGGCGGAATTCAATGTGATGATGAACCTGTGTGGCCTTCGTGTCCGCGGTGCT 505

QY 182 aIleSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGl 199
Db 506 CATGGGCTTCTCGAAACTGGGCAAGTCTCTTGAGATGCTGATCCTGGGCCGCTTCATCAT 565

QY 199 uGlyPheGlyValGlyIleIleSerTyThrValProValTyIleAlaGluIleSerPr 219
Db 566 CGGTGTGTACTCGGCCTGACCACAGGCTTCGTGCCCATGTATGTGGGTGAAGTGTCAAC 625

QY 219 oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh 239
Db 626 CACAGCCTTTCTGGGGCCCTGGGCACCCTGCACCAGCTGGGCATCGTCGTCGGCATCCT 685

QY 239 eLeuAlaTyLeuLeuGlyMetPheIle-----ProTrpArgLeuLe 253
Db 686 CATCGCCCGAGGTGTCTGGCCTTGACTCCATCATGGGCAACAAGGACCTGTGGCCCTGCT 745
```

```

QY 253 uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGl 273
Db 746 GCTGAGCATCATTTTCATCCCGCCCTGCTGCAGTGCATCGTGTGCCCTTCTGCCCGGA 805

QY 273 uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe 291
Db 806 GAGTCCCGCTTCTCTGCTC---ATCAACCGCAACGAGGAGAACCGGGCCCAAGAGTGTGCT 862

QY 291 uGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAspIleLysArgAl 311
Db 863 AAAGAGCTGCGCGGG---ACAGCTGACGTGACCATGACCTGCAGGAGATGAAGGAAGA 919

QY 311 aValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeu---AsnGlnLysLy 330
Db 920 GAGTCGGCAGATGATGCGGGAGAGAAGGTCACCATCTGGAGCTGTTCGGCTCCCCCGC 979

QY 330 sTyArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIl 350
Db 980 CTACCGCCAGCCCATCTCATCGCTGTGGTGTGCAGCTGTCCAGCAGCTGTCTGGCAT 1039

QY 350 eAsnGlyValLeuPheTyAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs 370
Db 1040 CAACGCTGTCTTCTATTACTCCACGAGCATCTTCGAGAACGGCGGGGTGCAGCAGCCTGT 1099

QY 370 pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe 390
Db 1100 GTATGCCACCATTTGGCTCCGGTATCGTCAACACAGGCCCTTCACCTGTGTGCTGTTTGT 1159

QY 390 uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuC 410
Db 1160 GGTGGAGCAGCAGGCGCGGACCCCTGCACCTCATAGGCCCTCGCTGGCATGGCGGTTG 1219

QY 410 sLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe 430
Db 1220 TGCCATACTCATGACCATCGCGCTAGCACTGTGGAGCAGCTACCCCTGG----- 1268

QY 430 rTyTrpIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe 450
Db 1269 -----ATGTCCTATCTGAGCATCGTGGCCATCTTTGGCTTTGTGGCCTTCTTTGA 1318

QY 450 rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValserIleLy 470
Db 1319 AGTGGGTCTCTGGCCCAICCCCATGGTTTCATCGTGGCTGAACCTTCAGCCAGGGTCCACG 1378

QY 470 sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe 490
Db 1379 TCCAGCTGCCATTGGCGGTGCAGGCTTCTCCAACTGGACCTCAAATTTTCATTTGGGCAT 1438

QY 490 tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValVa 510
Db 1439 GTGCTTCCAGTATGTGGAGCAACTGTGTGGTCCCTACGTCTTCATCATCTTCACCTGTGCT 1498

QY 510 lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly 527
Db 1499 CTTGGTTCTGTCTTCATCTTCACCTACTTCAAAGTTCTCTGAGACTAAAGGC 1550

RESULT 15
US-09-466-396A-135
; Sequence 135, Application US/09466396A
; Patent No. 6696247
; GENERAL INFORMATION:
; APPLICANT: Wang, TongLong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
```

US-09-466-396A-135				
Alignment Scores:				
Pred. No.:	7.68e-47	Length:	2856	
Score:	523.50	Matches:	151	
Percent Similarity:	45.91%	Conservative:	96	
Best Local Similarity:	28.07%	Mismatches:	229	
Query Match:	18.84%	Indels:	62	
DB:	4	Gaps:	12	
US-10-051-909-36 (1-553) x US-09-466-396A-135 (1-2856)				
Qy	27	ProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlySer-AsnArgGl	46	
Db	14	CCGAGTGAGCACGCCAGGAGCAGGAGACCAACGACGGGGTTCGAGTCAGAGTCGCAG	73	
Qy	46	yGlyAla-----GlyAlaGlyGluGluSerGlySerAspHisAspGlyValLeuAr	63	
Db	74	TGGGAGTCCCGGACCGGAGCAGGAGCCTGAGCGGGAGAGCGCGCTCGCACGCCCGTCG	133	
Qy	63	gArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAl	83	
Db	134	CCACCCGGGTACCCGG-----CGCAGCCAGAGCCACCAGCGCAGCGCTGC	178	
Qy	83	aProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysTh	103	
Db	179	CATGGAGCCCGCAGCAGCAAGAAGCTGACGGGTGCG-----CTCATGCTGGGTGT	226	
Qy	103	rLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProTh	123	
Db	227	GGGAGGAGCAGTGCTTGGCTCCCTGCGAGTTTGGCTACAACTGCGAGTCATCAATGCCCC	286	
Qy	123	rGlnAspAlaMetValArgAspLeuAen-----	132	
Db	287	CCAGAAGGTGATCGAGGAGTTCTACAACAGACATGGGTCCACCGCTATGGGAGAGCAT	346	
Qy	133	-----LeuSerIleSerGluPheSerAlaPheGl	142	
Db	347	CCTGCCACCAACCGCTCACCACGCTCTGGTCCCTCTCAGTGGCCATCTTTTCT-----	398	
Qy	142	ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy	162	
Db	399	-----GTTGGGGGCATGATGGCTCCTCTCTGTGGGCGCTTTTCGTTAACCG	445	
Qy	162	rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl	182	
Db	446	CTTTGGCCGGCGGAATTCAATGCTGATGATGAACCTGCTGGCCTTCGTGTCGCGCGTGCT	505	
Qy	182	aiIeSerPheAla-----LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGl	199	
Db	506	CATGGGCTTTCGAAACTGGGCAAGTCCTTTGAGATGCTGATCCTGGGCGCGTTTCATCAT	565	
Qy	199	uGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPr	219	
Db	566	CGGTGTACTGCGGCCTGACCACAGCGCTTCGTGCCCATGTATGTGGGTGAAGTGCACC	625	
Qy	219	oGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePh	239	
Db	626	CACAGCCTTTCGTGGGGCCCTGGGCACCCCTGCACCAGCTGGGCATCGTCGCGCATCCT	685	
Qy	239	eLeuAlaTyrLeuLeuGlyMetPheIle-----ProTrpArgLeuLe	253	
Db	686	CATCGCCCGAGGTTCGGCCTGGACTCCATCATGGGCAACAAGGACCTGTGGCCCTGTCT	745	
Qy	253	uAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGl	273	
Db	746	GCTGAGCATCATCTTCATCCCGGCCCTGCTGCAGTGCATCGTGTGCCCTTCTGCCCCGA	805	
Qy	273	uSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp-----CysGluThrSerLe	291	
Db	806	GAGTCCCGCTTCCTGCTC---ATCAACCGCAACGAGGAGAACCGGGCCAAGAGTGTGCT	862	
Qy	291	uGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAl	311	

Db	863	AAAGAAAGCTGCGCGGG---ACAGCTGACGTGACCCCATGACCTGCAGGAGATGAAGGAAGA	919
Qy	311	avalAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeu---AsnGlnLysLy	330
Db	920	GAGTCGGCAGATGCGGGAGAGAAGGTCACCATCCTGGAGCTGTTCGCTCCCCCGC	979
Qy	330	sTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIl	350
Db	980	CTACCGCCAGCCCATCCTCATCGCTGTGGTGTCTGCAGCTGTCCAGCAGCTGTCTGGCAT	1039
Qy	350	eAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAs	370
Db	1040	CAACGCTGTCTTCTATTACTCCAGGAGCATCTTCGAGAAGGCGGGGTGCAGCAGCCTGT	1099
Qy	370	pLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLe	390
Db	1100	GTATGCCACCATTTGGCTCCGGTATCGTCAACACGGCCCTTCACTGTCTGTCGCTGTTGT	1159
Qy	390	uLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCy	410
Db	1160	GGTGAGCGAGCAGGCGCGGACCCCTGCACCTCATAGGCCTCGCTGGCATGGCGGGTTG	1219
Qy	410	sLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSe	430
Db	1220	TGCCATACTCATGACCATCGCGCTAGCACTGCTGGAGCAGCTACCCCTGG-----	1268
Qy	430	rTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSe	450
Db	1269	-----ATGTCTATCTGAGCATCGTGGCCATCTTTGGCTTTGTGGCCTTCTTTGA	1318
Qy	450	rPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLy	470
Db	1319	AGTGGTCTCTGGCCCCATCCCATGGTTTCATCGTGGTGTAACCTTTCAGCCAGGGTCCACG	1378
Qy	470	sSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMe	490
Db	1379	TCCAGCTGCCATTGCCGTTGCAAGGCTTCTCCAACCTGGACCTCAAAATTTTCATTGTGGGCAT	1438
Qy	490	tThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValVa	510
Db	1439	GTGCTTCCAGTATGTGGAGCAACTGTGTGGTCCCTACGTCCTTCATCATCTTCACTGTGCT	1498
Qy	510	lSerAlaPheThrIleValPheValValLeuTrpValProGluThrLysGly	527
Db	1499	CCTGGTTCTGTTCTTTCATCTTCACCTACTTCAAAGTTCTCTGAGACTAAAGGC	1550

Search completed: April 13, 2005, 22:00:59  
Job time : 264.162 secs

This page blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 18:06:15 ; Search time 702.592 Seconds  
(without alignments)  
4775.020 Million cell updates/sec

Title: US-10-051-909-36  
Perfect score: 2779  
Sequence: 1 PSSSSSFRPAGKKKKKQG.....TIFVSLSIQRQLQWLPECLS 553

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074039\_14141/app\_query\_fasta\_1.1678  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10051909 @CGN 1 1 953 @runat 13042005 074039\_14141  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:
- 18: /cgn2\_6/ptodata/2/pubpna/US10F\_PUBCOMB.seq:
- 19: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:
- 20: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:
- 22: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2777	99.9	1953	13	US-10-051-909-35	Sequence 35, Appl
2	2133	76.8	2003	18	US-10-437-963-61061	Sequence 61061, A
3	2129.5	76.6	1797	17	US-10-425-114-26596	Sequence 26596, A
4	2104	75.7	1875	17	US-10-425-114-26550	Sequence 26550, A
5	1806.5	65.0	2049	17	US-10-425-114-24968	Sequence 24968, A
6	1806.5	65.0	2244	18	US-10-425-115-138759	Sequence 138759,
7	1797.5	64.7	1459	18	US-10-425-115-35618	Sequence 35618, A
8	1718	61.8	2202	17	US-10-424-599-132422	Sequence 132422,
9	1374	49.4	1304	18	US-10-437-963-63465	Sequence 63465, A
10	1356	48.8	1167	18	US-10-425-115-47714	Sequence 47714, A
11	1312.5	47.2	1105	18	US-10-425-115-47715	Sequence 47715, A
12	1221.5	44.0	957	17	US-10-425-114-28778	Sequence 28778, A
13	1185	42.6	1283	17	US-10-424-599-131066	Sequence 131066,
14	1182.5	42.6	1077	18	US-10-767-795-565	Sequence 565, App
15	1139.5	41.0	2105	17	US-10-424-599-132426	Sequence 132426,
16	1139	41.0	1027	17	US-10-425-114-5566	Sequence 5566, Ap
17	1081	38.9	1498	18	US-10-425-115-175050	Sequence 175050,
18	946.5	34.1	990	18	US-10-767-795-566	Sequence 566, App
19	944.5	34.0	1828	18	US-10-739-930-1250	Sequence 1250, Ap
20	941.5	33.9	785	18	US-10-425-115-140613	Sequence 140613,
21	920.5	33.1	1473	9	US-09-938-842A-2331	Sequence 2331, Ap
22	920.5	33.1	1473	11	US-09-938-842A-2331	Sequence 2331, Ap
23	909.5	32.7	1785	18	US-10-437-963-64291	Sequence 64291, A
24	899	32.3	1359	9	US-09-938-842A-1627	Sequence 1627, Ap
25	899	32.3	1359	11	US-09-938-842A-1627	Sequence 1627, Ap
26	893	32.1	1768	17	US-10-310-154-357	Sequence 357, App
27	888.5	32.0	1963	18	US-10-739-930-1789	Sequence 1789, Ap
28	877.5	31.6	1407	9	US-09-938-842A-1291	Sequence 1291, Ap
29	877.5	31.6	1407	11	US-09-938-842A-1291	Sequence 1291, Ap
30	777	28.0	1374	9	US-09-938-842A-1020	Sequence 1020, Ap
31	777	28.0	1374	11	US-09-938-842A-1020	Sequence 1020, Ap
32	759	27.3	1635	18	US-10-437-963-4124	Sequence 4124, Ap
33	733	26.4	1231	17	US-10-424-599-132423	Sequence 132423,
34	694	25.0	1319	18	US-10-425-115-39803	Sequence 39803, A
35	693.5	25.0	1779	17	US-10-424-599-104365	Sequence 104365,
36	679.5	24.5	778	17	US-10-425-114-22347	Sequence 22347, A
37	669.5	24.1	2080	16	US-10-168-651-28	Sequence 28, Appl
38	666	24.0	1258	17	US-10-424-599-132425	Sequence 132425,
39	663	23.9	1473	9	US-09-938-842A-1669	Sequence 1669, Ap
40	663	23.9	1473	11	US-09-938-842A-1669	Sequence 1669, Ap
41	658.5	23.7	1862	17	US-10-115-831-137	Sequence 137, App
42	657.5	23.7	600	18	US-10-021-323-14004	Sequence 14004, A
43	635	22.8	1242	18	US-10-425-115-177046	Sequence 177046,
44	626	22.5	668	18	US-10-425-115-137175	Sequence 137175,
45	625.5	22.5	890	17	US-10-425-114-127	Sequence 127, App

ALIGNMENTS

RESULT 1  
US-10-051-909-35  
; Sequence 35, Application US/10051909  
; Publication No. US20020199217A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Helentjaris, Tim  
; APPLICANT: Hitz, Bill  
; APPLICANT: Kinney, Tony  
; APPLICANT: Tingey, Scott  
; TITLE OF INVENTION: Plant Sugar Transport Proteins  
; FILE REFERENCE: BB1163 US CIP  
; CURRENT APPLICATION NUMBER: US/10/051,909  
; CURRENT FILING DATE: 2002-01-17  
; PRIOR APPLICATION NUMBER: 60/083,044  
; PRIOR FILING DATE: April 24, 1998  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 35  
; LENGTH: 1953











; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-227-G11\_FLI
US-10-425-114-24968

Alignment Scores:
Pred. No.: 2.75e-170 Length: 2049
Score: 1806.50 Matches: 366
Percent Similarity: 81.04% Conservative: 70
Best Local Similarity: 68.03% Mismatches: 79
Query Match: 65.01% Indels: 26
DB: 17 Gaps: 4

US-10-051-909-36 (1-553) x US-10-425-114-24968 (1-2049)
QY 34 LeuArgThrArgValMetGlyGlyGlySerAsnArgGlyGlyAlaGlyAlaGlyGluGlu 53
Db 314 CTCAGGCATGAGCTTCCGGACCCAGGAGAGTGGCGGGGAGGATG-GGGGCAGGACG-TCC 371
QY 54 SerGlySerAspHisAspGlyValLeuArgArgPProLeuLeuAsnThrGlySerTrpTyr 73
Db 372 TCCGCCTCCGAC-----CTGCGAAGCCGCTCATCAACACCCGGGAGCTGGTAC 419
QY 74 Arg-----MetSerSerArgGlnSerSerPheAlaProGlyThr 86
Db 420 CGCATGCCCGCGCGGTGGCGTGATGGCTCGCGGCAATCCAGCCTCATFGAGCGCTTG 479
QY 87 SerSerMetAlaVal---LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIle 105
Db 480 GGCTCCTCTCGGTTCTCTCTCCGCGACGTCGCTATCTCGGCCACGCTCTGCACGCTTATT 539
QY 106 ValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProThrGlnAsp 125
Db 540 GTCGCGTAGTCCCATCCAGTTCGGTTTCACATCGCGCTACTCTCGCCACGACGAGAC 599
QY 126 AlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSer 145
Db 600 GCCATCATTGCTGATCTCGGCCTCTCCCTCTCTGAGTTCTCCCTCTTCGGTTTCATTATCT 659
QY 146 AsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArg 165
Db 660 AATGTAGGGCGGATGGTAGCGCCATCTCCAGTGGGCAACTTGCAGAGTATATCGGCCGC 719
QY 166 LysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPhe 185
Db 720 AAGGGGTCTCTCATGATCGTCCGATTCCAAACATAATTGGGTGGCTCGCGATATATTC 779
QY 186 AlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIle 205
Db 780 GCAAAAGATTCTCTTCTTGTATGGTCCGCTGCTAGAGGATTGGAGTCGGTGTA 839
QY 206 IleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla 225
Db 840 ATATCGTATACAGTACCGGTTTATATTGCAGAAATCGCTCCTCAGGATCAGAGGGGAGCT 899
QY 226 LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGly 245
Db 900 CTTGGTTCTGTCAATCAGCTCTCCGTCACGATTGGTATATTGCTTGCTACTTGTTCGC 959
QY 246 MetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle 265
Db 960 ATGTTGTTCCCTGGAGAAATCTTGCTGTTCTAGGCATTTTACCCTTGTTCATCCTGATT 1019
QY 266 ProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGlu 285
Db 1020 CCTGGACTGTTCTTTGTGCTGAATCCCAAGGTGGTTGGCAAAATGGGGAAGATGGAG 1079
QY 286 AspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGluVal 305
Db 1080 GATTTTGAATATTTCATTGCAAGTCTCTGAGAGGATTTCAGACAGATATCACAGCAGAAGTA 1139
QY 306 AsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGlu 325

Db 1140 AATGAATAAAGAGATCATTTAGCATCATCGAGGAGGAGGACAAACCATAAGGTTCGCTGAT 1199
QY 326 LeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGln 345
Db 1200 ATCAAAACAGAAGAGATACAGTGTTCCTTGTGTATAGGAATCGGTCTCCTTGTCTGCAG 1259
QY 346 AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly 365
Db 1260 CAGCTAAGTGGTGTCAATGGCATCTATTTTATGCTGCGAGCATCTTCAAAAGCTGCTGGT 1319
QY 366 ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly 385
Db 1320 ATTACAAACAGTAATCTAGCAACATTTGGTTTAGGGGCTGTTCAGGTGATTGCTACTGGA 1379
QY 386 ValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSer 405
Db 1380 GTGACAAACCTGGTTGACTGACAAAGCTGGTCGAAGGCTTCTTCTCATTTATCCACCACA 1439
QY 406 GlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSer 425
Db 1440 GGAATGGTCATTACTCTTGTATTGTTTCTGTGTCAATTTTGTGAAGGACAAACATAGCT 1499
QY 426 GlnAspSerAsnSerTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe 445
Db 1500 GCTGGTTCCGACTTATCTCTGTAATGAGTATGCTTTCACCTGGCTGGACTTGTGGCATTT 1559
QY 446 ValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeu 465
Db 1560 GTGATTGCATTTTCTTGGCTTGGGAGCGGATTCGGTGGATCATTTATGCTGAGATCCTT 1619
QY 466 ProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSer 485
Db 1620 CCGTTTAACATCAAGAGCCTTGTCTGGAAGTGTTCGACCCCTGGCGAACTGGCTGACAGCA 1679
QY 486 PheAlaIleThrMetThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeu 505
Db 1680 TGGGCCATTACAATGACGGCAAGCCTGATGTTGAACCTGGAGCAGTGGAGGAACATTTGCT 1739
QY 506 SerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThr 525
Db 1740 ATCTACGCCGTCGTGTCTACCATGGCCCTCATTTTCGTGTGCTTGTGGTGCCTGAGACC 1799
QY 526 LysGly\*\*\*AsnSerArgGlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeu 545
Db 1800 AAGGGAAG-AACGCTAGAGGAAATCGCCTTCTCATTCCTGACACGTCGTCATGATCTA 1858
QY 546 -GlnTrp-----LeuProGluCysLeu 552
Db 1859 GGTATGGGAAGCCACACACCATGTGTCTTTCAGTCTGCCCGCGGTGCGTT 1910

RESULT 6
US-10-425-115-138759
; Sequence 138759, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 138759
; LENGTH: 2244
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577\_58031C.1
US-10-425-115-138759



Alignment Scores:			
Pred. No.:	3.15e-170	Length:	2244
Score:	1806.50	Matches:	366
Percent Similarity:	81.04%	Conservative:	70
Best Local Similarity:	68.03%	Mismatches:	79
Query Match:	65.01%	Indels:	26
DB:	18	Gaps:	4
US-10-051-909-36 (1-553) x US-10-425-115-138759 (1-2244)			
QY	34	LeuArgThrArgValMetGlyGlyGlySerAsnArgGlyGlyAlaGlyAlaGlyGluGlu	53
Db	359	CTCAGGCATGAGCTTCGGGACCAGGAGAGTGGCGGGAGGATG-GGGCAGGACG-TCC	416
QY	54	SerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySerTrpTyr	73
Db	417	TCCGCCTCCGAC-----CTGCGGAAGCCGCTCATCAACACCGGAGCTGGTAC	464
QY	74	Arg-----MetSerSerArgGlnSerSerPheAlaProGlyThr	86
Db	465	CGCATGCCGCGCGGGTGGCGTGATGGGCTCGCGCAATCTAGCCTCATGGAGCGATTG	524
QY	87	SerSerMetAlaVal---LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIle	105
Db	525	GGCTCCTCTCGGTTCTCTCCGCGACGTCGCTATCTCGCGACGCTCTGCACGCTCATT	584
QY	106	ValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProThrGlnAsp	125
Db	585	GTCGCGCTAGGTCCCATCCAGTTCGGTTTCACATCGCGGTACTCCTCGCCCCACGCAGGAC	644
QY	126	AlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSer	145
Db	645	GCCATCATTGCTGATCTCGGCCTCTCCCTCTCTGAGTTCTCCCTCTTCGGTTTCATTATCT	704
QY	146	AsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArg	165
Db	705	AATGTAGGGCGATGTTAGGCGCCATCTCCAGTGGGCAACTTGCAGAGTATATCGGCCGC	764
QY	166	LysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPhe	185
Db	765	AAGGGGTCTCTCATGATCGCTGGATTCCAAACATAATTGGGTGGCTCGCGATATCATT	824
QY	186	AlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIle	205
Db	825	GCAAAAGATTCTCTTCTTGTTTATGGGTGGGTGCTAGAAGGATTTGGAGTCGGTGTA	884
QY	206	IleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla	225
Db	885	ATATCGTATACAGTACCGGTTTATATTGCAGAAATCGCTCCTCAGGATCAGAGGGAGCT	944
QY	226	LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGly	245
Db	945	CTTGGTTCTGTCAATCAGCTCTCCGTACAGATTGGTATATTGCTTGCCTACTTGTGGC	1004
QY	246	MetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle	265
Db	1005	ATGTTTGTCCCTGGAGAATTCTTGCTGTTCTAGGCATTTTACCTTGTTCATCCTGATT	1064
QY	266	ProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGlu	285
Db	1065	CCTGGACTGTTCTTTGTGCCTGAATCCCAAGGTGGTTGGCAAAATGGGAAGATGGAG	1124
QY	286	AspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluVal	305
Db	1125	GATTTTGAATATTCTATTGCAAGTTCTGAGAGGATTTTCAGACAGATATCACAGCAAGTA	1184
QY	306	AsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGlu	325
Db	1185	AATGAATAAAGAGATCATTAGCATCATCGAGGAGGAGGACCAACATAAGGTTTCGCTGAT	1244
QY	326	LeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGln	345
		:::      :::      :::      :::      :::      :::      :::      :::	

QY	346	AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly	365
Db	1305	CAGCTAAGTGGTGTCAATGCGATTCTATTTTATGCTGCGAGCATCTTCAAAGCTGCTGCT	1364
QY	366	ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly	385
Db	1365	ATTACAAACAGTAATCTAGCAACATTTGGTTTAGGGGCTGTTCAGGTGATTGCTACTGA	1424
QY	386	ValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleSerThrSer	405
Db	1425	GTGACAAACCTGGTTGACTGACAAAGCTGGTCGAAGGCTTCTTCTCATTATTCCACCACA	1484
QY	406	GlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsnIleSer	425
Db	1485	GGAATGGTGCATTACTCTTGTTATTGTTTCTGTGTCAATTTTGTGAAGGACAACATAGCT	1544
QY	426	GlnAspSerAsnSerTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe	445
Db	1545	GCTGGTTCGCACCTTATCTCTGTATGATGATGCTTCTCACTGGCTGGACTTGTGGCATTT	1604
QY	446	ValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeu	465
Db	1605	GTGATTGCATTTTCTTGGCTTGGAGCGGATTCCTGGATCATTTATGTCTGAGATCCTT	1664
QY	466	ProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSer	485
Db	1665	CCTGTTAACATCAAGAGCCTTGTGGAAGTGTTCGACCTTGGCGAACTGGCTGACAGCA	1724
QY	486	PheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeu	505
Db	1725	TGGGCCATTACAATGACGCAAGCCTGATGTTGAACCTGGAGCAGTGGAGGAACCTTTGCT	1784
QY	506	SerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThr	525
Db	1785	ATCTACGCCGTCGTGCTACCATGGCCCTCATTTTCGTGCTTGTGGTGCCTGAGACC	1844
QY	526	LysGly***AsnSerArgGlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeu	545
Db	1845	AAGGGAAG-AACGCTAGAGGAAATCGCCTTCTCATTCGCTGACACGCTCGTCATGATCTA	1903
QY	546	-GlnTrp-----LeuProGluCysLeu	552
Db	1904	GGTATGGGAAGCCACACACCATGTGTCTATTGAGTCTGCCGCGGTGCGTT	1955
RESULT 7			
US-10-425-115-35618			
; Sequence 35618, Application US/10425115			
; Publication No. US20040214272A1			
; GENERAL INFORMATION:			
; APPLICANT: La Rosa, Thomas J.			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated wi			
; TITLE OF INVENTION: Plants			
; FILE REFERENCE: 38-21(53222)B			
; CURRENT APPLICATION NUMBER: US/10/425,115			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 369326			
; SEQ ID NO 35618			
; LENGTH: 1459			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; OTHER INFORMATION: Clone ID: MRT4577_132481C.1			
US-10-425-115-35618			
Alignment Scores:			
Pred. No.:	1.32e-169	Length:	1459
Score:	1797.50	Matches:	372
Percent Similarity:	98.15%	Conservative:	0

Db	1245	ATCAAAACAGAAGAGATACAGTGTTCCTCCCTTGTGATAGGAATCGGTCTCCTTGTCTCTGCAG	1304						
Qy	346	AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly	365						
Db	1305	CAGCTAAGTGGTGTCAATGGCATTCTATTTATGCTGCGAGCATCTTCAAAGCTGCTGGT	1364						
Qy	366	ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly	385						
Db	1365	ATTACAAACAGTAATCTAGCAACATTGGTTTAGGGGCTGTTCAAGTGATTGCTACTGGA	1424						
Qy	386	ValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSer	405						
Db	1425	GTGACAACTGGTTGACTGACAAAGCTGGTCGAAGGCTTCTTCTCATTTATTCACCACCA	1484						
Qy	406	GlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSer	425						
Db	1485	GGAATGGTCATTACTCTTGTATTGTTTCTGTGTCAATTTTGTGAAGGACAACATAGCT	1544						
Qy	426	GlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe	445						
Db	1545	GCTGGTTCGCACCTTATACTCTGTAATGAGTATGCTTTCACCTGGCTGGACTTGTGGCATT	1604						
Qy	446	ValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeu	465						
Db	1605	GTGATTGCATTTTCTCTTGGCTGGAGCGATTCCGTGGATCATTTATGCTGAGATCCTT	1664						
Qy	466	ProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSer	485						
Db	1665	CCTGTTAACATCAAGAGCCTTGCTGGAAGTGTTCGACCCCTGGCGAACTGGCTGACAGCA	1724						
Qy	486	PheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeu	505						
Db	1725	TGGGCCATTACAATGACGGCAAGCCTGATGTTGAACCTGGAGCAGTGGAGGAACCTTTGCT	1784						
Qy	506	SerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpValProGluThr	525						
Db	1785	ATCTACGCCGTCGTGTCTACCATGGCCCTCATTTTCGTGTGCTTGTGGGTGCCTGAGACC	1844						
Qy	526	LysGly***AsnSerArgGlyAspThrIlePheValSerLeuSerIleGlnArgGlnLeu	545						
Db	1845	AAGGGAAG-AACGCTAGAGGAATCGCCTTCTCATTCGCTGACACGTCGTATGATCTA	1903						
Qy	546	-GlnTrp-----LeuProGluCysLeu 552							
Db	1904	GGTATGGGAAGCCACACACCATGTGTGTCATTGAGTCTGCCCGGTGCGTT 1955							
RESULT 7									
US-10-425-115-35618									
; Sequence 35618, Application US/10425115									
; Publication No. US20040214272A1									
; GENERAL INFORMATION:									
; APPLICANT: La Rosa, Thomas J.									
; APPLICANT: Kovalic, David K.									
; APPLICANT: Zhou, Yihua									
; APPLICANT: Cao, Yongwei									
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With									
; TITLE OF INVENTION: Plants									
; FILE REFERENCE: 38-21(53222)B									
; CURRENT APPLICATION NUMBER: US/10/425,115									
; CURRENT FILING DATE: 2003-04-28									
; NUMBER OF SEQ ID NOS: 369326									
; SEQ ID NO 35618									
; LENGTH: 1459									
; TYPE: DNA									
; ORGANISM: Zea mays									
; FEATURE:									
; OTHER INFORMATION: Clone ID: MRF4577_132481C.1									
US-10-425-115-35618									
Alignment Scores:									
Pred. No.:	1.32e-169	Length:	1459						
Score:	1797.50	Matches:	372						
Percent Similarity:	98.15%	Conservative:	0						

Best Local Similarity: 98.15%		Mismatches: 7	
Query Match: 64.68%		Indels: 3	
DB: 18		Gaps: 0	
US-10-051-909-36 (1-553) x US-10-425-115-35618 (1-1459)			
QY	175	ProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMet	194
Db	3	CCCAAATTCATCGGTTGGCTTGCATCTCCTTTGCAAAAGA-GCCTCATTTCTATATATG	61
QY	195	GlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProValTyrIle	214
Db	62	GGACGATTGCTTGAAGGGTTTGGTGTGGGATCATATCCTACACGGTACCGGTATACATA	121
QY	215	AlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerVal	234
Db	122	GCAGAGATATCTCCTCAGAACATGAGGGAGCTCTTGGTTCTGTGAACCAAGTTGCTGTG	181
QY	235	ThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAla	254
Db	182	ACCTTTGGCATATTCTTGGCCTATTGTCTGGCATGTTTATTCTCTTGAGACTTCTTGCT	241
QY	255	ValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSer	274
Db	242	GTGATTGGAGCCTTGCCCTGCACAATGTTGATTCTCTGGCTATTCTTCATTCAGAACT	301
QY	275	ProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeu	294
Db	302	CCCAGATGGCTGGCAAAGATGAATTTGATGGGAAGATTGTGAGACGTCCCTACAAGTGCTG	361
QY	295	ArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSer	314
Db	362	AGGGGGTTGGGACTGACATCAACAACAGAAAGTGAATGATATAAAGAGGGCAGTGGCATCA	421
QY	315	SerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrPro	334
Db	422	TCAAGTA--AGGACCAACAATCAGTTTTCAAGAAATTAACCAAAAGAAATACCGCACGCCA	479
QY	335	LeuLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeu	354
Db	480	CTACTTCTAGGATTTGGCCTACTTGTACTGCAAAATCTTAGTGAATCAACGGTGTACTG	539
QY	355	PheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCys	374
Db	540	TTTTTATGCAAGTAGCATCTTCAAAGCTGCAGGGGTTACAAACAGCGACTTGGCCACCTGT	599
QY	375	SerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAla	394
Db	600	TCACTTGGTGCTATTTCAGGTCTCTTGCTACTGGAGTTACAACATGGCTGTAGACCGAGCT	659
QY	395	GlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaVal	414
Db	660	GGACGACGCATCCTTCTCATTTATTTCTACCTCTGGCATGACTCTATGCCCTTCTGCCGTT	719
QY	415	SerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeu	434
Db	720	TCTGTTGTATTTTTTCTCAAGGATAACATTTACAGGATTTCTAACTCATACTACATCTTA	779
QY	435	ThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGly	454
Db	780	ACAAATGATCTCCCTTGTGGTATTGTGTCTTTTGTGTCATTACCTTCTCGTTTGGTATGGGT	839
QY	455	AlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGly	474
Db	840	GCCATTCCATGGCTCATGATGTCTGAGATCCTCCCGGTTAGCATCAAGAGCCTTGCGCGA	899
QY	475	SerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeu	494
Db	900	AGCATCGCAACACTGGCCAACTGGCTGACATCCTTCGCCATAACAATGACGACGAACCTG	959
QY	495	MetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSerAlaPheThr	514
Db	960	ATGCTCACGTGGAGTGTGGAGGCACCTTTTCTCTCGTACATGGTTGTGAGCGCCCTCACC	1019

QY	515	IleValPheValValLeuTrpValProGluThrLysGly***AsnSerArgGlyAspThr	534
Db	1020	ATCGTTTTTGTGTCTCTTGGTGCCGGAGACGAAGGGAG-AACTCTAGAGGAGATACA	1078
QY	535	IlePheValSerLeuSerIleGlnArgGlnLeuGlnTrpLeuProGluCysLeuSer	553
Db	1079	ATTTTCGTTTCGCTGAGCAATTCAGCGTTCAGCTGCAATGGTTGCCGAGTGTATCT	1135
RESULT 8			
US-10-424-599-132422			
; Sequence 132422, Application US/10424599			
; Publication No. US20040031072A1			
; GENERAL INFORMATION:			
; APPLICANT: La Rosa Thomas J			
; APPLICANT: Kovalic David K			
; APPLICANT: Zhou Yihua			
; APPLICANT: Cao Yongwei			
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With			
; TITLE OF INVENTION: plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(53223)B			
; CURRENT APPLICATION NUMBER: US/10/424,599			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 285684			
; SEQ ID NO 132422			
; LENGTH: 2202			
; TYPE: DNA			
; ORGANISM: Glycine max			
; FEATURE:			
; NAME/KEY: unsure			
; LOCATION: (1)..(2202)			
; OTHER INFORMATION: unsure at all n locations			
; FEATURE:			
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1			
US-10-424-599-132422			

Alignment Scores:		2.23e-161		Length: 2202	
Pred. No.:		1718.00		Matches: 334	
Score:		83.67%		Conservative: 76	
Percent Similarity:		68.16%		Mismatches: 76	
Best Local Similarity:		61.82%		Indels: 5	
Query Match:		17		Gaps: 2	
DB:					

US-10-051-909-36 (1-553) x US-10-424-599-132422 (1-2202)

QY	52	GluGluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeuAsnThrGlySer	71
Db	255	GAAGAGAGTGGCGATGCGAGGAT-----CTTCAGAAACCGTTCCTCCACACAGGGAGT	308
QY	72	TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal	91
Db	309	TGTTACAAGATGGGTTCCAGGCAGTCCAGCATCATG-----GGATCCTCCACTCATGTT	362
QY	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle	111
Db	363	ATCCGCGACGCGCGCGTTTCCGTCTCTTCTCGCTCATCGTCGCTTGGTCCCCATT	422
QY	112	GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu	131
Db	423	CAATTGGCTTCACGTGTGGGTATTCTTCTCCAACCCAAAGGGGTATAGTTCGCGATCTA	482
QY	132	AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyMetVal	151
Db	483	AACCTCTCTATTTCGGAGTTTCTTCTTTTGGATCTTTGTCTAATGTGGAGCAATGGTG	542
QY	152	GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle	171
Db	543	GGAGCTATAGCTAGTGGTCAGATAGCTGAATACATCGGCGCAAGGGTCATTGATGATT	602
QY	172	AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe	191
Db	603	GCTGCGATCCCCAATATAATAGGGTGGCTTGCTATTCTTTTGCCAAAGATTCTCTCGTTT	662

```
Qy 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211
|||
Db 663 TTGTATATGGGAGGTTGTTGGAAGGTTTGGCGTTGGGATTATCTTTATGTGTGCCT 722

Qy 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231
|||
Db 723 GTTTATATAGCTGAGATTGCACCTCAAAACTTGAGAGGTGGCCTTGGGTCAGTGAACCG 782

Qy 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251
|||
Db 783 CTCCTGTGTACAATTGGCATTATGCTGGCTTATCTGTTGGTCTTTTGTCAACTGGAGA 842

Qy 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIle 271
:::
Db 843 GTCCTTGCAATTCTAGGAATTTGCCTTGTACAGTATTAATACCTGGATTATTTTCATA 902

Qy 272 ProGluSerProArgTyrLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu 291
|||
Db 903 CCTGAATCCCCAGATGGTTGGCCCAAGATGGGATGATAGATGAGTTTGAGACTTCTTTG 962

Qy 292 GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla 311
|||
Db 963 CAAGTGTTACGAGGATTTGACACTGATATATCTGTTGAAGTACATGAATAATTAAGAGATCT 1022

Qy 312 ValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr 331
|||
Db 1023 GTGGCTTCAACGGGAAAGAGCTGCAATCCGATTTCAGATCTCAAGAGGAAAGATAT 1082

Qy 332 ArgThrProLeuLeuGlyIleGlyLeuValLeuGlnAsnLeuSerGlyIleAsn 351
|||
Db 1083 TGGTTCCCGTTAATGTTGGTATTGGATTACTTGTCTTCAGCAATTATCTGGTATCAAT 1142

Qy 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371
|||
Db 1143 GGAATTTGTTCTATTCAACTACCATCTTTGCAAAATGCAGGAATTCATCCAGCGAAGCT 1202

Qy 372 AlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGlyValThrTrpLeuLeu 391
|||
Db 1203 GCTACAGTTGGACTTGGAGCCGTTTCAAGTCATAGCAACTGGAATTTCCACATGTTGGTG 1262

Qy 392 AspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeu 411
|||
Db 1263 GACAAAAGTGGCGGAGGCTGCTTCTTAATAATATCTCATCTGTAATGACAGTTAGCCTT 1322

Qy 412 LeuAlaValSerValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyr 431
|||
Db 1323 CTCATTGTTTCTATAGCATTTTATCTGGAGGGGTTGTATCAGAGGATTACATTTATTTC 1382

Qy 432 TyrIleLeuThrMetIleSerLeuValGlyIleValSerPheValIleThrPheSerPhe 451
|||
Db 1383 AGCATTTGGGAATAGTTTCTATTGTTGGACTCGTGGCTATGGTGATGGGTTCTCTCTA 1442

Qy 452 GlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeuProValSerIleLysSer 471
|||
Db 1443 GGTCTGGGACCCATCCCTTGGCTTATAATGTCTGAGATACTTCCAGTGAATATAAAGGC 1502

Qy 472 LeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSerPheAlaIleThrMetThr 491
|||
Db 1503 CTTGTGCGAGCATAGCGACAATGGGAAATTTGGCTGATTTTCGTGGGGGATCACGATGACT 1562

Qy 492 ThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPheLeuSerTyrMetValValSer 511
|||
Db 1563 GCTAACTTGCTTTTGAATTGGAGCAGTGGAGGGACATTTACAATCTACACAGTCGTAGCT 1622

Qy 512 AlaPheThrIleValPheValValLeuTrpValProGluThrLysGly**AsnSerArg 531
|||
Db 1623 GCCTTTACTATAGCTTTTATAGCAATGTGGGTTCTCTGAGACCAAGGAAG-AACATTGGA 1681

Qy 532 GlyAspThrIlePheValSerLeuSerIle 541
|||
Db 1682 AGAAATTTCAGTTTTCCTTCAGATAGATATA 1711
```

RESULT 9

```
US-10-437-963-63465
; Sequence 63465, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 63465
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64704C.1
US-10-437-963-63465

Alignment Scores:
Pred. No.: 2,89e-127 Length: 1304
Score: 1374.00 Matches: 293
Percent Similarity: 68.89% Conservative: 48
Best Local Similarity: 59.19% Mismatches: 58
Query Match: 49.44% Indels: 97
DB: 18 Gaps: 6

US-10-051-909-36 (1-553) x US-10-437-963-63465 (1-1304)
```

```
Qy 51 GlyGluGluSerGlySerAspHisAspGly-----ValLeuArgArg 64
|||
Db 72 GGCAGAGAGAGCGCGGCGAGGATGGGGGCGGACGGCGTCGGCTCGGACCTGCGGAAG 131

Qy 65 ProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAlaPro 84
|||
Db 132 CCGTCTCTCCACACGCGGGAGCTGCTACAAGATGTCTGCGCGGGCGGCGGGGATG 191

Qy 85 GlyThr-----SerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeu 101
|||
Db 192 GGGTCGGCCCTCGGCTCTCCGCTACTCCTCCGCGACTCCTCCGCTCTCCGCGTCTC 251

Qy 102 CysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSer 121
|||
Db 252 TGCACCTCATCGTCGCCCTCGGCCCATCCAGTTCGGCTTACCTGCGGCTTCTCCTCG 311

Qy 122 ProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPhe 141
|||
Db 312 CCCACCCAGGACGCCCATCATCTCCGACCTCGGCCTCACCCCTCTCCGAGTTCTCGCTG 371

Qy 142 GlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGlu 161
|||
Db 372 GGGTCGCTGTGGAACGTCGGGGCAATGGTGGCGCCCATCGCCAGCGGCCAGATCGCCGAG 431

Qy 162 TyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeu 181
|||
Db 432 TACATCGGCCCGCAAGGGGTCTCTCATGATCGCGGCAATTCTTAACATTATTGGGTG 491

Qy 182 AlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlyPhe 201
|||
Db 492 GCAATATCTTTTGCAAGGACTCA----- 515

Qy 202 GlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsn 221
|||
Db 516 -----TGAGTGCCAGTTTATATAGCAGAAATTGCTCCACAAACA 554
```





Db 681 GAGATATCTCCACAGAACATGAGAGGGCTCTTGGCTCTGTGAACCAAGTTGTCTGTAAACC 740  
Qy 236 PheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLeuAlaVal 255  
Db 741 CTTGGTATCATGTTTGCTTATTTGCTCGGCTGTGTTGTTCTTGGAGGCTTCTTGCAATA 800  
Qy 256 IleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerPro 275  
Db 801 ATAGGAACCTTGCCCTGCATAGTTGTGTATACCTGGCCTTTTCTTCAATCCGGAATCTCCA 860  
Qy 276 ArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg 295  
Db 861 AGATGGCTGGCAAGGATGAATATGATGGATGATTCGAGACTTCTCTACAAGTTCTGAGA 920  
Qy 296 GlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSer 315  
Db 921 GGATTGATCTGTACATCACTGCAGAAATTGAATGATATAAGAGAGCAGTAATGTCAGCA 980  
Qy 316 SerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyr----- 331  
Db 981 AACAAAAGGGCTACAATTTCGTTTCCAGAGTTGAACCAAAAAAAGGCGGC 1040  
Qy 332 ArgThrProLeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsn 351  
Db 1041 CGTACGCCCCCTAAATTCTAGGAATGGGCCTACTTGTGTTGCAGCAGCTGAGTGGGATAAAC 1100  
Qy 352 GlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeu 371  
Db 1101 GGTATAATCTTTATGCAGGTAGCATCTT-AAAGCTGCAAGGCTCAAGAACAGTAACCTG 1159  
Qy 372 AlaThr 373  
Db 1160 GACACA 1165

RESULT 11  
US-10-425-115-47715  
; Sequence 47715, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 47715  
; LENGTH: 1105  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_143523C.1  
US-10-425-115-47715

Alignment Scores:  
Pred. No.: 3.26e-121 Length: 1105  
Score: 1312.50 Matches: 271  
Percent Similarity: 92.31% Conservative: 5  
Best Local Similarity: 90.64% Mismatches: 10  
Query Match: 47.23% Indels: 16  
DB: 18 Gaps: 1  
  
US-10-051-909-36 (1-553) x US-10-425-115-47715 (1-1105)  
  
Qy 1 ProSerSerSerSerPheArgProAlaGlyLysLysLysLysLysLysAsnGlnGly 20  
Db 212 CCTTCCTCTCGTCTCTCTTCAGGCCAGCGGCAAGAAGAAGA--AAAAATCAAGGC 269  
Qy 21 LeuArgArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGly 40

Db 270 TTGCGGCGAGAGGCTGTGCCCCGGCCGACCGCGAGCGAACTTCGTACGCGCGTCATGGGT 329  
Qy 41 GlyGlySerAsnArgGlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAspGly 60  
Db 330 GGCGGCAGCAACAGAGGCGCGCGCGCGCGCGCGAGAGAGCGGAGCGGACACGACGGT 389  
Qy 61 ValLeuArgArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSer 80  
Db 390 GTGCTGGGAGGCCGCTGCTCAACACGGGGAGCTGGTACCGGATGAGCTCGGGCAGTCC 449  
Qy 81 SerPheAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPhe 100  
Db 450 AGCTTTGCCCCGGGGACCTCCTCCAG-GCCGCTCTGCGCGAGTCCACAGTCTCCGCCCTC 508  
Qy 101 LeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSer 120  
Db 509 CTCTGCAGCTCATCGTCGCGCTCGGCCCATCCAGTTCGGCTTCACCAGCGGCTTCTCC 568  
Qy 121 SerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAla 140  
Db 569 TCCCCGACCCAGGACGCCATGGTTCGGGACCTCAACCTCTCTATCTCCGAGTTCTCGGCG 628  
Qy 141 PheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAla 160  
Db 629 TTCGGATCGCTGTCCAACGTCGGCGGCATGGTCGGGCGATCGCGAGCGGCGAGATGGCC 688  
Qy 161 GluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrp 180  
Db 689 GAGTACATTGGCCGTAAAGGGTCGTTGATGATTGCTGCAATCCCAAATATCATCGGTTGG 748  
Qy 181 LeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGly 200  
Db 749 CTTGCGATCTCCTTTGCAAAAGATGCCCTCATTTCTGTATATGGGACGCTTGCTTGAAGGA 808  
Qy 201 PheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGln 220  
Db 809 TTCGGTGTCCGTATCATTTCCCTATGTGGTACCGGTATACATAGCAGAGATATCTCCACAA 868  
Qy 221 AsnMetArgGlyAlaLeuGlySerValAsnGlnLeu----- 232  
Db 869 AACATGAGAGGGGCTCTTGGCTCTGTGAACCAGGT-TTGCTTATACTTGCAAGCATCTCA 927  
Qy 233 -----SerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe 247  
Db 928 TCTATTGTGAGTTATCTGTAAACCCTTGGTATCATGTTGCCTATTGCTCGGCTTGTTT 987  
Qy 248 IleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly 267  
Db 988 GTTCCTTGGAGGCTTCTTGCAGTAATAGGAACCTTGCCCTGCATAGTGTGTATACCTGGC 1047  
Qy 268 LeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAsp 286  
Db 1048 CTTTCTTCTCATTCGGAATCTCCAAGATGGCTGGCAAGATGAATATGATGGATGAT 1104  
  
RESULT 12  
US-10-425-114-28778  
; Sequence 28778, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 28778

```

; LENGTH: 957
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4764-010-D11_FLI
US-10-425-114-28778

Alignment Scores:
Pred. No.: 3.41e-112 Length: 957
Score: 1221.50 Matches: 249
Percent Similarity: 85.27% Conservative: 23
Best Local Similarity: 78.06% Mismatches: 33
Query Match: 43.95% Indels: 14
DB: 17 Gaps: 6

US-10-051-909-36 (1-553) x US-10-425-114-28778 (1-957)
QY 23 ArgGluAlaValProGlyArgPro-----AlaSerGluLeuArgThrArgValMetGly 40
Db 2 CGGGAACGGTCCCAACCCCGGAGAAACGAAACCAAAATCTCGCTCGAGCTG 61
QY 41 GlyGlySerAsn-ArgGlyGly-----AlaGlyAlaGlyGlu---G1 53
Db 62 GCGCGGAGGCGCGCGCGCCATGGCGAGGAGCAGGACGTCGGGGCGGCGAGTACGA 121
QY 53 uSerGlySerAspHisAspGlyValLeuArgArgProLeuLeu---AsnThrGlySerTr 72
Db 122 GAGCGGAGCGACCAAGACGGCGCGCTGCAGAGCGCGTGTGCCCCAACAGCGGGAGCTG 181
QY 72 pTyrArg-----MetSerSerArgGlnSerSerPhe--AlaProGlyThrSerSerMe 89
Db 182 GTACCGGATGGGATGGGGTCGGCCAGTCCAGCCTCAACGCCGCCGACCTCTCTCCAT 241
QY 89 tAlaValLeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuG1 109
Db 242 GCGCGTCGCGCGAGTCCCACTCCGCTTCGCGCTTCTGTGCACGCTCATCGTCGCGCTCGG 301
QY 109 yProIleGlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValAr 129
Db 302 CCCCATCCAGTTCGGCTTCACCGCGGCTACTCTCTCCCGCAGCGCAGCCAGCATCCG 361
QY 129 gAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyG1 149
Db 362 GGACCTCAACCTCTCCATCTCCGAGTCTCGGTGTTTCGGCTCGCTGCCAACGTGGCGC 421
QY 149 yMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLe 169
Db 422 CATGGTCGGGGCGATCGCCAGCGCGCAGATGGCCGAGTACGTGTGGCGCAAGGGTCGT 481
QY 169 uMetIleAlaAlaIleProAsnIleIleGlyTyrLeuAlaIleSerPheAlaLysAspAl 189
Db 482 GATGATTGCAGCAATTCCAAACGTCAATGGTTGGCTTGGCTCTCCTTTGCAAAAGACTC 541
QY 189 aSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrTh 209
Db 542 TTCATTTCTGTATATGGACGCTTGCTTGAAGGATTCGGTGTCCGTATCAITTCCTATGT 601
QY 209 rValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerVa 229
Db 602 GGTACCGGTATACATAGCAGAGATATCTCCACAAACATGAGAGGGGCTCTTGGCTCTGT 661
QY 229 lAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIlePr 249
Db 662 GAACCAAGTTATCTGTAACCTTGGTATCATGTTTGGCTACTTGTCTCGGCTTGTTGTTCC 721
QY 249 oTyrArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPh 269
Db 722 TTGGAGGCTTCTTCAGTAATAGGAACCTTGCCCTGCATAGTGTGTATACCTTGGCTTTT 781
QY 269 ePheIleProGluSerProArgTyrPheAlaLysMetAsnLeuThrGluAspCysGluTh 289
Db 782 CTTTCATTCCGGAATCTCCAAAGATGGCTGGCAAAAGATGAATATGATGATGATGCGAGAC 841
```

```

289 rSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLy 309
Db 842 TTCTCTACAAGTTCTGAGAGGATTCGATGCTGACATCACTGCAGAATTGAATGATATAA 901
QY 309 sArgAlaValAlaSerSerSerLysArgThrThrIleSerPheGlnGluLeuAsn 327
Db 902 GAGAGCAGTAATGTCAGCAACAAAGGGCTACAATTGTTTCCAAGAGTTGAAC 956

RESULT 13
US-10-424-599-131066
; Sequence 131066, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 131066
; LENGTH: 1283
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_8935C.1
US-10-424-599-131066

Alignment Scores:
Pred. No.: 2.39e-108 Length: 1283
Score: 1185.00 Matches: 223
Percent Similarity: 81.84% Conservative: 61
Best Local Similarity: 64.27% Mismatches: 51
Query Match: 42.64% Indels: 12
DB: 17 Gaps: 2

US-10-051-909-36 (1-553) x US-10-424-599-131066 (1-1283)
QY 201 PheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGln 220
Db 2 TTTGGGGTAGGAATAATATCTTACACGGTGCCTGTGTACATAGCTGAGATATCACCTCCA 61
QY 221 AsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeu 240
Db 62 AACTTGAGGGGGGGTTTGGTTTCAGTTTAACCAAGCTCTCTGTCCACCATTTGAATTATGCTG 121
QY 241 AlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeuPro 260
Db 122 GCATATCTGCTGGGATTTTGTGTGAATGGAGAATCCTTGCAATATAGGAATTTTACCC 181
QY 261 CysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTyrPheLeuAlaLys 280
Db 182 TGTACAATATTGATACCTGGCCTATTTTTCATTTCCCGAGTCCCTTAGATGGCTGGCAAAA 241
QY 281 MetAsnLeuThrGluAspCysGluThrSerSerLeuGlnValLeuArgGlyPheGluThrAsp 300
Db 242 ATGGGAATGACAGAGAATTTGAAACTTCTCTTGCAAGTGTCTCCGAGGCTTTGAGACTGAT 301
QY 301 IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThr 320
Db 302 ATTTCTGTTGAAGTGAATGAAATTAAGAGGGCTGTTGCTTCAACAAACAGAAAGACTACA 361
QY 321 IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGly 340
Db 362 GTTCGATTGACAGACCTCAACAAAGAGATATGGCTTCCCTTAATGATTTGGAATTGGA 421
QY 341 LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIle 360
Db 422 CTGCTTATTTTGCAACACAGCTTCTGGAATTAATGGTGTCTTTTTTATTCAGTACCATC 481
```





Db 1018 AGTACCTTTGGAGTTGGCTTCATTCAGGTCAATGCTACTGCTTTAAACCACATGGTTG 1075

RESULT 15

US-10-424-599-132426

; Sequence 132426, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 132426

; LENGTH: 2105

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(2105)

; OTHER INFORMATION: unsure at all n locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_90588C.1

US-10-424-599-132426

Alignment Scores:

Pred. No.:	1.79e-103	Length:	2105
Score:	1139.50	Matches:	228
Percent Similarity:	70.55%	Conservative:	69
Best Local Similarity:	54.16%	Mismatches:	58
Query Match:	41.00%	Indels:	67
DB:	17	Gaps:	2

US-10-051-909-36 (1-553) x US-10-424-599-132426 (1-2105)

QY 187 LysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIle 206

Db 320 CAGGATTCCTCATTTTGTATATGGGGAGCCTGCTGGAAGGTTTTTGGCGTCGGGATTAGC 379

QY 207 SerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeu 226

Db 380 TCTTATAGGTGCCTGTTTATATAGCTGAGATTGCACCTCAAACTTGAGAGGTGGCCYT 439

QY 227 GlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMet 246

Db 440 GGATCAGTGAACCAAGCTCTCTATTACTATTGGCAATATGCTGGCTTATCTTTTGGGCCTT 499

QY 247 PheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIlePro 266

Db 500 TTTGTCAACTGGAGAGTGTGCAATTTAGGAATTTTGCCCTGTACAGTATTAAATACCT 559

QY 267 GlyLeuPhePheIleProGluSerProArgTrpLeu----- 278

Db 560 GGATTATTTTTCATACCTGAATCCCCCAGATGGTTGGTATGGATATGTTATGCATTATAT 619

QY 278 ----- 278

Db 620 TGATCAAGTTTGTACTTTTTCATCTTTGTTAAATAVACTGTTTCCTTTCAGAAAAAGGATG 679

QY 278 ----- 278

Db 680 GTAACCTTTTCTACTCTTGGTAATTATCTTTCTGCTTCTCTTTTCATTTTATTATTITA 739

QY 279 -----AlaLysMetAsnLeuThrGluAspCy 287

Db 740 TTTTTCCTGTTGGTGGGAATCGGGGAAGTGTAGGCTAAGATGGGGATGACAGATGAGTT 799

QY 287 sGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAs 307

Db 800 TGAGACTTCTTTTGCAAGTGTTACGAGGATTTGACACAGATATATCTGTTGAAGTATATGA 859

QY 307 PileLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGluLeuAs 327

Db 860 AATTAAAGAGATCTGTGGCTTCAACGGGAAAAAGAGCTACAATCCGATTTCAGATCTCAA 919

QY 327 nGlnLysLysTyrArgThrProLeuLeu-----LeuGlyIleGlyLe 341

Db 920 GAGGAAAAGATATTGGTTCCCTTAATGGTATAATATGACACTGTAGTTGGTATTGGATT 979

QY 341 uLeuValIleuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePh 361

Db 980 ACTTGTCTTCAGCAGTTATCTGGTATCAATGGAGTTTGTCTTATTCAACTACCATCTT 1039

QY 361 eLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnVa 381

Db 1040 TGCAAAATGCAGGAATTTCCGTCACGGAAGCTGCTACTGTTGGACTTGGTGCCGTTCA 1099

QY 381 lLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIl 401

Db 1100 CATAGCAACTGGAATTTCCACGCTGTTGGTGGACAAAAAGTGGTCGGAGGCTGCTCTA 1159

QY 401 eIleSerThrSerGlyMetThrLeuCysLeuAlaValSerValValPhePheLeuLy 421

Db 1160 GATATCCTCATCTGTAATGACTGTAGCCTTCTCATTTGTTTCAATAGCATTTTATCTGGA 1219

QY 421 sAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGl 441

Db 1220 GGGGGTTGTATCTGAGGATTCAATCTATTTCAGCATGTTGGGAATAGTTTCTGTTGTGG 1279

QY 441 YIleValSerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMe 461

Db 1280 ACTTGTGTTATGTTGTCATTTGGGTTCTCTTGGTCTGGGACCCCATCCCTTGGCTTATA 1339

QY 461 tSerGluIleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAs 481

Db 1340 GTCTGAGATACTTCCAGTGAATATAAAGGGCCTTGTCTGGCAGCATAGCGACAATGGG 1399

QY 481 nTrpLeuThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGl 501

Db 1400 TTGGCTGATTTTCGTGGGTGATCACGATGACCGCTAACTTACTTTTGAATTGGAACAGTGG 1459

QY 501 YGlyThrPheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTr 521

Db 1460 AGGGACATTTTACAATCTACACAGTGGTAGTGCCTTTTACTATTGCTTTTATAGCATTATG 1519

QY 521 pValProGluThrLysGly\*\*AsnSerArgGlyAspThrIlePheValSerLeuSerIl 541

Db 1520 GGTTCCTGAGACCAAGGGAAG-AACATTGGAAGAAATTCAGTTTCTTTCAGATAGATGT 1578

QY 541 e 541

Db 1579 T 1579

Search completed: April 14, 2005, 01:49:24

Job time : 731.592 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2005, 12:44:19 ; Search time 3817.87 Seconds  
(without alignments)  
5513.431 Million cell updates/sec

Title: US-10-051-909-36  
Perfect score: 2779  
Sequence: 1 PSSSSSFRPAGKKKKKNQG.....TIFVSLSIQRQLQWLPECLS 553

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10051909/runat\_13042005\_074037\_14060/app\_query.fasta\_1.1678  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10051909 @CGN 1 1 6628 @runat\_13042005\_074037\_14060 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_htc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1733.5	62.4	1837	3	CNSOACUC
2	1219	43.9	854	7	CN125032
3	1102	39.7	868	6	CB675064
4	1065	38.3	682	6	CA191028
5	1021	36.7	635	6	CA143053
6	1004.5	36.1	764	6	CB648359
7	984.5	35.4	1155	7	CK211005
8	984	35.4	674	1	AV939950
9	976	35.1	714	6	CA262285
					BX815012 Arabidops
					CN125032 RH0H1_8 A
					CB675064 OSJNE101
					CA191028 SCCRT2C0
					CA143053 SCQSRT203
					CB648359 OSJNE111
					CK211005 FGAS02283
					AV939950 AV939950
					CA262285 SCEQLB201

10	975	35.1	681	1	AV913420	AV913420
11	962.5	34.6	718	7	CR291514	CR291514
12	962.5	34.6	762	1	AJ794429	AJ794429
13	932	33.5	701	6	CD871121	CD871121
14	930.5	33.5	1675	3	CNS0A8HU	AX819759 Arabidops
15	925	33.3	769	4	BI933646	BI933646 EST553535
16	919.5	33.1	1737	3	CNS09ZVC	BX832192 Arabidops
17	914	32.9	772	6	CB683315	CB683315 OSJNE111
18	907.5	32.7	1666	3	CNS0A8HV	BX819760 Arabidops
19	907	32.6	666	5	BQ766951	BQ766951 EBR008_SQ
20	904.5	32.5	627	6	CA180355	CA180355 SCCCST300
21	904	32.5	652	6	CA201877	CA201877 SCRFFL102
22	902	32.5	716	7	CF451024	CF451024 EST687369
23	901.5	32.4	748	6	CB347677	CB347677 CAB2SG000
24	896.5	32.3	1601	3	CNS0A620	BX822241 Arabidops
25	896	32.2	685	5	BQ862009	BQ862009 QGC1f12.Y
26	893	32.1	566	4	BM325827	BM325827 PIC1_53_C
27	890	32.0	791	6	CB892803	CB892803 EST645595
28	887.5	31.9	746	6	CD483180	CD483180 atr01-31m
29	882	31.7	594	6	CD203945	CD203945 HS1_2_E01
30	882	31.7	772	7	CF835377	CF835377 UCRCS03_0
31	870.5	31.3	744	1	AJ796408	AJ796408 AJ796408
32	869	31.3	641	6	CA246696	CA246696 SCSGFL5C0
33	864.5	31.1	927	7	CK153511	CK153511 FGAS03213
34	859	30.9	548	6	CA143695	CA143695 SCRURT200
35	850	30.6	627	4	BJ249193	BJ249193 BJ249193
36	847	30.5	663	7	CN906216	CN906216 010929ABD
37	845	30.4	670	5	BQ855360	BQ855360 QGB25P02.
38	838	30.2	595	2	BE599181	BE599181 P11_86_D0
39	830.5	29.9	651	4	BG522368	BG522368 20_64_Ste
40	828	29.8	667	6	CD893209	CD893209 G118.123C
41	827	29.8	626	7	CN904481	CN904481 010831ABD
42	825	29.7	656	5	BQ913239	BQ913239 QHA6L20.Y
43	822	29.6	1054	3	AY111571	AY111571 Zea mays
44	821	29.5	686	7	CO116723	CO116723 GR_Eb019
45	820.5	29.5	714	6	CA262252	CA262252 SCEQLB201

ALIGNMENTS

RESULT 1

CNSOACUC

LOCUS

DEFINITION

CNSOACUC 1837 bp mRNA linear HTC 06-FEB-2004  
Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
GSLTLS24ZB10 of Adult vegetative tissue of strain col-0 of  
Arabidopsis thaliana (thale cress).

ACCESSION

VERSION BX815012.1 GI:42472886

KEYWORDS

SOURCE HTC; GSLT cDNA.

ORGANISM

Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,  
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,  
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.

TITLE

Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished

REFERENCE 2 (bases 1 to 1837)

AUTHORS

TITLE Direct Submission

JOURNAL

Submitted (18-NOV-2003) Genoscope - Centre National de Sequençage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

COMMENT

The sequences are based on single pass reads.  
Life technologies (a division of Invitrogen) members carried out  
full-length libraries construction : Temple G.  
Genoscope members carried out sequencing and annotation : Castelli  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,





Sun,F., Sullivan,R., Lim,S., Eastman,A. and Pratt,L.H.  
An EST database from Sorghum: acid- and alkaline-treated roots  
Unpublished (2003)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
the Human Genome Center, University of Tokyo Institute of Medical  
Science; plant material and RNA prepared at Texas A & M University;  
sequencing done in the Laboratory for Genomics and Bioinformatics,  
University of Georgia. Sequence ends have been trimmed to exclude  
vector and regions below Phred quality 16. Three-prime sequences  
are presented as their reverse complement and have been trimmed to  
exclude polyA.  
Seq primer: Sugs (CTTCTGCTCTAAAGCTGCG).  
FEATURES  
source  
1..854  
/organism="Sorghum bicolor"  
/mol\_type="mRNA"  
/cultivar="BTx623"  
/db\_xref="taxon:4558"  
/clone="RHO1 8 A07 A002"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Acid- and alkaline-treated roots"  
/note="Organ: Root; Vector: pME18S-FL3; Site 1: XhoI;  
Site 2: XhoI; The library was prepared from PolyA+ RNA  
from 8-day-old roots harvested from BTx623 sorghum  
seedlings grown in hydroponic culture. HCl was added to a  
pH of 3.0 to some seedlings, KOH to a pH of 9.0 for  
others. Roots were harvested 3, 12 and 27 hr after  
addition and pooled for RNA preparation. Double-stranded  
cDNA was cloned unidirectionally into different DraIII  
sites of the pME18S-FL3 vector (5-prime DraIII site is  
CACTGTGTG, 3-prime DraIII site is CACCATGTG). XhoI excises  
the cDNA insert."  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.38e-115 Length: 854  
Score: 1219.00 Matches: 250  
Percent Similarity: 93.43% Conservative: 6  
Best Local Similarity: 91.24% Mismatches: 11  
Query Match: 43.86% Indels: 7  
DB: 7 Gaps: 3  
US-10-051-909-36 (1-553) x CN125032 (1-854)  
Qy 3 SerSerSerPheArgProAlaGlyLysLysLysLysLysAsnGlnGlyLeuArg 22  
Db 51 TCTTCTTCCTCCTTCAGGCCAGCGGC---AAGAAGAAGAGAAAAATCAAAGCTTGCGG 107  
Qy 23 -ArgGluAlaValProGlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyG1 42  
Db 108 GCGAGAGCGCGTGCCCGCGCAAGCGAGCGAA-----GCCATGGGTGGCGCGG 158  
Qy 42 ySerAsnArgGlyGlyAlaGlyAlaGlyGluGluSerGlySerAspHisAspGlyValle 62  
Db 159 CAGCAACAGAGGC-----GGCGCCAGCAGGAGAGCGGCAGCAGCAGCGCGGGCT 212  
Qy 62 uArgArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPh 82  
Db 213 GCGGAAGCCGCTGCTCCACACGGGCAGCTGTACCGTATGAGCTCGGCGCAGTCCAGCGT 272  
Qy 82 eAlaProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCy 102  
Db 273 CGCCCCCGGGCCTCCTCCATGGCCGTCGTGGGGAGTCCCACGTCCTCGCCTTCTCTCG 332  
Qy 102 sThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerPr 122

Db 333 CACACTCATCGTCGCGCTCGGCCCCCATTCAGTTCGGCTTCACCAGCGGCTTCTCCTCCCC 392  
Qy 122 oThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheG1 142  
Db 393 GACCCAGGACACCATGGTCCGGACCTCAACCTCTCCATCTCCGAGTTCTCGCGGTTCGG 452  
Qy 142 ySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTy 162  
Db 453 CTCGCTGTCCAACGTCGGCGCCATGGTCGGGGCGATCGCCAGCGGCAGATGGCCGAGCA 512  
Qy 162 rIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAl 182  
Db 513 CATTGGCCGTAAAGGGTTGTTGATGATTGCTGCAATCCCGAATATCATCGGTGGCTTGC 572  
Qy 182 aIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheG1 202  
Db 573 GATCTCCTTTGCAAAAGATGCTCATTTCTTTATATGGGACGATTGCTTGAAGGGTTTGG 632  
Qy 202 yValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMe 222  
Db 633 TGTCCGCATCATATCTATACGTACCTGTATACATAGCAGAGATATCTCCTCAGAACAT 692  
Qy 222 tArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTy 242  
Db 693 GAGAGGAGCTCTTGGTTCTGTGAACCAAGTTGTCCGTGACCTTTGGCATATTCTTTGGCCTA 752  
Qy 242 rLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysth 262  
Db 753 TTTGCTCGGCATGTTTGTTCCTTGGAGACTTCTAGCTGTATCGGAGCCTTGGCCCTGCAC 812  
Qy 262 rMetLeuIleProGlyLeuPhePheIleProGluSerPro 275  
Db 813 AGTGTGATTCTCTGGACTATTCTTCATTCCAGAAATCTCCC 852  
RESULT 3  
CB675064  
LOCUS  
DEFINITION  
CB675064.1 868 bp mRNA linear EST 09-APR-2003  
OSJNEe10I13.f OSJNEe Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNEe10I13 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 868)  
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,  
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.  
Large-scale identification of ESTs involved in the interaction  
between rice and Magnaporthe grisea  
Unpublished (2003)  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ  
85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: gta aaa cga cgg cca gtg  
BACKWARD: gga aac agc tat gac cat g  
Plate: 10 row: 1 column: 13  
Seq primer: gta aaa cga cgg cca gtg.  
Location/Qualifiers  
1..868  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNEe10I13"

/tissue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNEe"  
/note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:  
XhoI; 24 hrs after inoculation with Rice Blast (70-15)"  
ORIGIN

Alignment Scores: 2.01e-103 Length: 868  
Pred. No.: 1102.00 Matches: 208  
Score: 87.89% Conservative: 46  
Percent Similarity: 71.97% Mismatches: 35  
Best Local Similarity: 39.65% Indels: 0  
Query Match: 6 Gaps: 0  
DB:

US-10-051-909-36 (1-553) x CB675064 (1-868)

QY	206	IleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAla	225
Db	1	ATATCCTATGTGGTGCCAGTTATATAGCAGAAATGCTCCACAAACAATGAGAGGCT	60
QY	226	LeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGly	245
Db	61	CTTGGTTCAGTCAATCAGCTTCTGTTACTATTGGCATATTGCTTGCTACTTGTTAGGC	120
QY	246	MetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIle	265
Db	121	ATGTTTGTCCCTGGAGAAATCTCTCCGTTCTCGGTATTTTACCATGTTCAATTCTGATA	180
QY	266	ProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGlu	285
Db	181	CCTGGATTGTTCTTTATCCCTGAATCACCAGGTGGCTGGCAAAAATGGGAAAGATGAAG	240
QY	286	AspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGluVal	305
Db	241	GATTTTGAATCCTCGCTGCAAGTATTGCGAGGATTGAAACAGACATCGCTGTAGAAGTA	300
QY	306	AsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGlnGlu	325
Db	301	AATGAATAAAGAGAACAGTTCATCAATCATCCGGAGGAGGACGACAATACGATTGCGAGAT	360
QY	326	LeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeuGln	345
Db	361	ATCAGCAGAGAAGATATAGTACCTCTAATGATAGGAATTGGTCTCCTGTACTGCAG	420
QY	346	AsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAlaGly	365
Db	421	CAGTTAAGTGGTGTGAATGGCATCTATTTTATGCTGCTAGTATCTTCAAAGCCGCCGGT	480
QY	366	ValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThrGly	385
Db	481	CTTACGAATAGTAATCTAGCAACATTTGGTCTGGGGGTGTTTCAGGTGGTGTACTGGA	540
QY	386	ValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeuIleIleSerThrSer	405
Db	541	GTGACAACTTGGTTGACTGACAAAGCTGGTCGACGACTCCTCCTCATTATCTCTACTACA	600
QY	406	GlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsnIleSer	425
Db	601	GGGATGACCATTACTCTTGTGTTGTTTCTGTGTCATTTTGTGAAGGACAACTAACT	660
QY	426	GlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSerPhe	445
Db	661	AATGGTTCTCATTTTACTCTGTAATGAGTATGCTTTCGCTGGTTCGGGCTTGTGGCATTT	720
QY	446	ValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIleLeu	465
Db	721	GTGATTTTCATTTTCTCTTGGTTTGGGAGCCATACCATGGATCATAAATGTCGAGATTCIT	780
QY	466	ProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThrSer	485
Db	781	CCTGTAATATCAAGAGCCCTTGCTGGAAGCGTTGCAACCCTTGCAAACTGCTGTGACGGCT	840

486 PheAlaIleThrMetThrThrAsnLeu 494  
:::|||||:::|  
841 TGGCTCATGATGACAGCAAGCTTG 867

RESULT 4  
CA191028  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CA191028 682 bp mRNA linear EST 24-SEP-2003  
SCCIRT2C03A08.g RT2 Saccharum officinarum cDNA clone SCCIRT2C03A08  
5', mRNA sequence.

CA191028  
CA191028.1 GI:35135776  
EST.  
Saccharum officinarum  
Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum  
complex.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 682)  
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.  
The libraries that made SUCEST  
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)  
Contact: Arruda P  
Centro de Biologia Molecular e Engenharia Genetica  
Universidade Estadual de Campinas  
Caixa Postal 6010, 13083-970, Campinas SP, Brazil  
Tel: 55 19 3788 1137  
Fax: 55 19 3788 1089  
Email: parruda@unicamp.br  
Clone distribution: clone distribution information can be found  
through the Brazilian Clone Collection Center (BCCC) at  
http://www.bcccenter.fcav.unesp.br  
Plate: C03 row: A column: 08  
Seq primer: T7 Promoter Primer.  
Location/Qualifiers

FEATURES  
source

1..682  
/organism="Saccharum officinarum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4547"  
/clone="SCCIRT2C03A08"  
/lab\_host="DH10B"  
/clone\_lib="RT2"  
/note="Organ: Root tips(0.3cm-long) from adult plants;  
Vector: pSport1; Site 1: SalI; Site 2: NotI; An  
unidirectional cDNA library generated from [Root  
tips(0.3cm-long) from adult plants]. cDNA was prepared  
from polyA+ mRNA using SuperScript Plasmid System Kit  
(Invitrogen). The double-strand cDNAs were fractionated  
in a sepharose CL-2B 40cm-columns and fragments sizing  
between 0.8 and 1.5 Kb were directionally cloned into the  
vector. Details of each source of RNA and library  
construction can be obtained at  
http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Alignment Scores: 9.58e-100 Length: 682  
Pred. No.: 1065.00 Matches: 218  
Score: 97.36% Conservative: 3  
Percent Similarity: 96.04% Mismatches: 6  
Best Local Similarity: 38.32% Indels: 1  
Query Match: 6 Gaps: 0  
DB:

US-10-051-909-36 (1-553) x CA191028 (1-682)

QY	256	IleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPheIleProGluSerPro	275
Db	2	ATCGGAGCCTTGCCTGCACAGTGTGTGATTCCTGGACTATTCTTCAGATCTCCC	61
QY	276	ArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArg	295
Db	62	AGATGGCTGGCAAGATGAATTTGATGGAAGATTGCGAGACGTCCTACAGTGTGAGG	121



Qy	296	GlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSer	315
Db	122	GGGTTTGAGACTGC-ATCACGACAGAAGTGAATGATATAAAGAGGCGAGTGACATCATCA	180
Qy	316	SerLysArgThrThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeu	335
Db	181	AGTAAGAGGACTACAATCAGTTTTCAGAATTAACCAAAAGAAATACCGCACGCCGCTA	240
Qy	336	LeuLeuGlyIleGlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPhe	355
Db	241	CTTCTAGGAATTGGCCTACTTGTACTGCAAAATCTAAGTGGAAATCAACGGTGTACTGTT	300
Qy	356	TyrAlaSerSerIlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSer	375
Db	301	TATGCAAGTAGCATCTTCAAGCTGCAGGTGTTACAAACAGCGACTTGGCCACCTGTTCA	360
Qy	376	LeuGlyAlaIleGlnValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGly	395
Db	361	CTTGGAGCTATCCAGTCTTGTCTACTGGAGTTACGACATGGTTGTTAGACAGAGCTGGA	420
Qy	396	ArgArgIleLeuLeuIleIleSerThrSerGlyMetThrLeuCysLeuAlaValSer	415
Db	421	CGACGCATGCTTCTCATATTCTTCTACCTCTGGCATGACTCTATGCCCTTCTTGCCGTTCT	480
Qy	416	ValValPhePheLeuLysAspAsnIleSerGlnAspSerAsnSerTyrTyrIleLeuThr	435
Db	481	GTTGTATTTTCTCAAGGATAAGATTTTCACAGGATTCTAACTCGTACTACATCTTAACT	540
Qy	436	MetIleSerLeuValGlyIleValSerPheValIleThrPheSerPheGlyMetGlyAla	455
Db	541	ATGATCTCCTTGGTTGCTATCGTGGCTTTTGTCAATTACCTTCTCGTTTGGTATGGGTGCC	600
Qy	456	IleProTrioMetMetSerGluIleLeuProValSerIleLysSerLeuGlyGlySer	475
Db	601	ATTCCATGGCTCATGATGTCTGAGATCCTTCCAGTTAGCATCAAGAGTCTCGGCGGAAGC	660
Qy	476	IleAlaThrLeuAlaAsnTirp	482
Db	661	ATCGCGACACTGGCCAAATGG	681

```

/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCQSRT2032A08"
/lab_host="DH10B"
/clone_lib="RT2"
/note="Organ: Root tips(0.3cm-long) from adult plants;
Vector: pSport1; Site 1: SalI; Site 2: NotI; An
unidirectional cDNA library generated from [Root
tips(0.3cm-long) from adult plants]. cDNA was prepared
from polyA+ mRNA using SuperScript Plasmid System Kit
(invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 Kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucest.lad.ic.unicamp.br/public"

```

# ORIGIN

```

Alignment Scores:
  Pred. No.:      3.18e-95
  Score:          1021.00
  Percent Similarity: 98.10%
  Best Local Similarity: 96.68%
  Query Match:    36.74%
  DB:             6
                Length: 635
                Matches: 204
                Conservative: 3
                Mismatches: 4
                Indels: 0
                Gaps: 0

US-10-051-909-36 (1-553) x CAL43053 (1-635)

```

```
RESULT 6
CB648359          CB648359          764 bp      mRNA      linear      EST 08-APR-2003
LOCUS             OSJNEb11L10.f OSJNEb Oryza sativa (japonica cultivar-group) cDNA
DEFINITION        clone OSJNEb11L10 5', mRNA sequence.
ACCESSION         CB648359
VERSION           CB648359.1 GI:29643352
KEYWORDS          EST.
SOURCE            Oryza sativa (japonica cultivar-group)
ORGANISM          Oryza sativa (japonica cultivar-group)
                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                  Ehrhartoideae; Oryzeae; Oryza.
REFERENCE         1 (bases 1 to 764)
AUTHORS           Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
                  Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE             Large-scale identification of ESTs involved in the interaction
                  between rice and Magnaporthe grisea
JOURNAL           Unpublished (2003)
COMMENT           Contact: Rod Wing
                  Arizona Genomics Institute
                  University of Arizona
                  Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
                  85721-0088, USA
                  Tel: 520 626 3967
                  Fax: 520 621 9288
                  Email: http://genome.arizona.edu
PCR PRIMERS      FORWARD: gta aaa cga cgg cca gtc
                  BACKWARD: gga aac agc tat gac cat g
                  Plate: 11 row: L column: 10
                  Seq primer: gta aaa cga cgg cca gtc.
FEATURES         location/Qualifiers
                  1..764
                  /organism="Oryza sativa (japonica cultivar-group)"
                  /mol_type="mRNA"
                  /cultivar="Nipponbare"
                  /db_xref="taxon:39947"
                  /clone="OSJNEb11L10"
                  /tissue_type="Leaf"
                  /dev_stage="3 week"
                  /lab_host="DH10B"
                  /clone_lib="OSJNEb"
                  /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
                  XhoI; 24 hrs after inoculation with Rice Blast (Che
                  86061)"
ORIGIN
Alignment Scores:
Pred. No.:        2.21e-93          Length:      764
Score:            1004.50           Matches:     202
Percent Similarity: 87.40%          Conservative: 13
Best Local Similarity: 82.11%       Mismatches:  24
Query Match:      36.15%           Indels:      7
DB:               6                Gaps:        2
US-10-051-909-36 (1-553) x CB648359 (1-764)
QY      28  GlyArgProAlaSerGluLeuArgThrArgValMetGlyGlyGlySerAsnArgGlyGly  47
Db      35  GGAGCCGAGCGGCGAGCGAGCGGGCGCGATGAACGGCGCGGGAGCAGCGCGGGAGGT  94
QY      48  AlaGlyAlaGlyGluGluSerGlySerAspHisAsp-----GlyValLeuArg  63
Db      95  GCGTCGTGCGCCGACGAGCGGCGAGCGGCGAGCGGCGGCGGCGGCGGCGGCTGCGG  154
QY      64  ArgProLeuLeuAsnThrGlySerTrpTyrArgMetSerSerArgGlnSerSerPheAla  83
Db      155 AAGCCGCTGCTGAACACGCGGGAGCTGGTACAGGATGGGATCGGGTCCAGCCTCGCGCC  214
QY      84  ProGlyThrSerSerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysThr  103
Db      215 -----TCCTCCATGGCCGCCCATCCGGAGTCCCACGCTCCGCGCTTCTCTGCACG  265
```

```
QY      104  LeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProThr  123
Db      266 CTCATCTGTCGGCTCGGCCCATCCAATTCCGATTCCAGCGGCTTCTCTCACCACC  325
QY      124  GlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySer  143
Db      326 CAGGACGCCATCATCCGCGACCTCAAGCTCTCCAATCCTCGGAGTTCTCGGCTTCGGGTCG  385
QY      144  LeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIle  163
Db      386 CTGTCCAACGTCGGCGCCATGTCGGAGCGATCGCCAGTGGGCAGATGGCGGAGTACATT  445
QY      164  GlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIle  183
Db      446 GGCCGGAAGGTCGTTGATAATTGGCGGGTTCCTTAACATCATTTGGTTGGCTTGCCATC  505
QY      184  SerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyVal  203
Db      506 TCCTTTGCAAAAGACGCGTCATTTTATACATGGGACGCTTGCTTGAAGGTTTTGGTGT  565
QY      204  GlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArg  223
Db      566 GGTGTTATATCATATACGGTGCCAGTATACATAGCAGAGATATCTCATCAGAACACAAGA  625
QY      224  GlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeu  243
Db      626 GGAGCACTTGCTCGTCCGTGAACCAAGTTGTCCGTTACCATTTGGTATCTTGTGGCCCTATT  685
QY      244  LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet  263
Db      686 CTAGGCATGTTGTTCCTCTGGAGGCTGCTTGCACTGATAGGAAGCATCCCATGTACATTG  745
QY      264  LeuIleProGlyLeuPhe  269
Db      746 TTAATACCTGGTCTATTTC  763
RESULT 7
CK211005          CK211005          1155 bp      mRNA      linear      EST 08-DEC-2003
LOCUS             FGAS022833 Triticum aestivum FGAS: Library 5 GATE 7 Triticum
DEFINITION        aestivum cDNA, mRNA sequence.
ACCESSION         CK211005
VERSION           CK211005.1 GI:39573395
KEYWORDS          EST.
SOURCE            Triticum aestivum (bread wheat)
ORGANISM          Triticum aestivum
                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
                  Poideae; Triticeae; Triticum.
REFERENCE         1 (bases 1 to 1155)
AUTHORS           Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
                  Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
                  Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,
                  Penniket,C., Roach,J.L. and Sarhan,F.
                  Functional Genomics of Abiotic Stress In Wheat and Canola Crops
                  Unpublished (2003)
                  Contact: Wm L Crosby
                  Bioinformatics
                  University of Saskatchewan, Department of Computer Science
                  1C101 Engineering Building, 57 Campus Drive, Saskatoon,
                  Saskatchewan, S7N 5A9, Canada
                  Tel: 306 966 1769
                  Fax: 306 966 2033
                  Email: fgas_ests@cs.usask.ca
                  This sequence is the direct result of the Base calling software
                  Phred (default parameters). It is the raw base calls. To aid in the
                  identification of the high quality insert the software Lucy
                  (default parameters) has been run on this sequence. Lucy identified
                  the region [44,808].
                  Plate: L5B024 row: L column: 10.
FEATURES         location/Qualifiers
                  1..1155
source
```

/organism="Triticum aestivum" /mol_type="mRNA" /db_xref="taxon:4565" /clone_lib="Triticum aestivum FGAS: Library 5 GATE 7" /note="Vector: pCMV.SPORT6; Crown and developmental stages of spike formation in wheat cultivar Norstar. 4 mRNA populations were combined before constructing the library. The first mRNA population is from lcm crown sections after 30 days of cold acclimation. The second is from lcm crown sections after 11 days of deacclimation (before deacclimation plants were fully vernalized for 49 days). The third is from different developmental stages of spike formation (5 to 50mm) that still have not emerged from the leaf (dissection required). The last is from different developmental stages of spike and seed formation after having emerged from the leaf (visible). First strand synthesis in this library was done in the presence of methylated dCTP thereby protecting from internal cleavage with NotI."									
ORIGIN									
Alignment Scores:									
Pred. No.:	5.09e-91	Length:	1155						
Score:	984.50	Matches:	195						
Percent Similarity:	82.83%	Conservative:	51						
Best Local Similarity:	65.66%	Mismatches:	51						
Query Match:	35.43%	Indels:	3						
DB:	7	Gaps:	0						
US-10-051-909-36 (1-553) x CK211005 (1-1155)									
QY	244	LeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMet	263						
Db	54	GTAGGCATGTTTGTTCCTTGGAGAATTCTCTCAGTCCTTGGCATTTTACCTTGCTCAATC	113						
QY	264	LeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeu	283						
Db	114	CTGATACCTGGTGTGTTCTTCATCCCGGAATCACCAAGGTGGCTGGCAAAATGGAAAG	173						
QY	284	ThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThr	303						
Db	174	ATGAGGATTTTGAATCTTCGCTGCAGGTTCTGCGGGATTTGAAACTGATATCTCAGCA	233						
QY	304	GluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThrIleSerPhe	323						
Db	234	GAAGTAAATGAAATAAAGAGATCAGTTGCCTCATCTAGGAGGAGGACAACCATACGATT	293						
QY	324	GlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuVal	343						
Db	294	GCAGAGATCAAAACACAGAGATATAGTGTTCCTCTTATGATAGGAATTGGTCTCCTTATA	353						
QY	344	LeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAla	363						
Db	354	CTTCAGCAACTAAGTGGTGTCAATGGCATCTTCTTCTATGTCTGCAAGTATCTTCAAAGCT	413						
QY	364	AlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAla	383						
Db	414	GCTGGTCTTAAAAATAGTAATCTAGCAACATGTGGTTTGGGGCTGTTTCAGGTGGTTGCT	473						
QY	384	ThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgGlyIleLeuLeuIleIleSer	403						
Db	474	ACTGGAATCACAACCTGTTGACTGACAAAGCTGGTCGACGGCTACTTCTCATTATCTCT	533						
QY	404	ThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePheLeuLysAspAsn	423						
Db	534	GCTGTAGGGATGACAGTCAGTCTTCTCGTTGTTTCTGTGTCATTTGTTGAAGGAAAC	593						
QY	424	IleSerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleVal	443						
Db	594	ATAGAGGAAGCTTCTCATTTACACTCTGTGATGAGTATGCTTTTCACTGGCTGGACTTGTG	653						
QY	444	SerPheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGlu	463						

Db	654	GCATTTGTGCTTGCAATTTCCCTTGGCATGGGAGCCCATCCCATGGATCATATAATGTCTGAG	713						
QY	464	IleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeu	483						
Db	714	ATCCTTCTCTGTTAAACATCAAGAGTCTGCCGGGAAGCACCCCCACCTTCGCAAACTGGATG	773						
QY	484	ThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThr	503						
Db	774	ACGTCCTGGCTCATCAGATGACCGGGAGCTTGATGCTAAACTGGAGCAACGGAGGAACC	833						
QY	504	PheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpValPro	523						
Db	834	TTGCTATA-TTCGCGGGGGTGTCCATGGGCACCCCTCTCTTCGTGTCCCTGTGCGTCCCG	892						
QY	524	GluThrLysGly**AsnSerArgGlyAspThrIlePheValSerLeuSer	540						
Db	893	GAGACCAAGGGAAG-AACATTGAGGAAA-TCGGTTTTCGTTCCGCTGAAC	941						
RESULT 8									
AV939950	AV939950 674 bp mRNA linear EST 18-JAN-2002								
LOCUS	AV939950 K. Sato unpublished cDNA library, strain H602 adult,								
DEFINITION	heading stage top three leaves Hordeum vulgare subsp. spontaneum								
	cDNA clone bah24ml2 5', mRNA sequence.								
ACCESSION	AV939950								
VERSION	AV939950.1 GI:18235747								
KEYWORDS	EST.								
SOURCE	Hordeum vulgare subsp. spontaneum								
ORGANISM	Hordeum vulgare subsp. spontaneum								
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;								
	Poideae; Triticeae; Hordeum.								
REFERENCE	1 (bases 1 to 674)								
AUTHORS	Sato,K., Saisho,D. and Takeda,K.								
TITLE	Barley EST sequencing project in NIG and Okayama Univ								
JOURNAL	Unpublished (2002)								
COMMENT	Contact: Tadasu Shin-i								
	Center For Genetic Resource Information								
	National Institute of Genetics								
	1111 Yata, Mishima, Shizuoka 411-8540, Japan								
	Tel: 81-559-81-6856								
	Fax: 81-559-81-6855								
	Email: tshini@genes.nig.ac.jp.								
FEATURES	Location/Qualifiers								
source	1..674								
	/organism="Hordeum vulgare subsp. spontaneum"								
	/mol_type="mRNA"								
	/strain="H602"								
	/sub_species="spontaneum"								
	/db_xref="taxon:77009"								
	/clone="bah24ml2"								
	/tissue_type="top three leaves"								
	/dev_stage="adult, heading stage"								
	/clone_lib="K. Sato unpublished cDNA library, strain H602								
	adult, heading stage top three leaves"								
ORIGIN	.								
Alignment Scores:									
Pred. No.:	2.44e-91	Length:	674						
Score:	984.00	Matches:	193						
Percent Similarity:	91.93%	Conservative:	12						
Best Local Similarity:	86.55%	Mismatches:	17						
Query Match:	35.41%	Indels:	1						
DB:	1	Gaps:	0						
US-10-051-909-36 (1-553) x AV939950 (1-674)									
QY	53	GlusSerGlySerAspHisAspGlyValLeuArgArgPro-LeuLeuAsnThrGlySerTr	72						
Db	3	GAGAGCGGAGCGGACCATGACACCGGNAAGCAGCCGCTGCTGGTCAACACGGGGAGCTG	62						
QY	72	pTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValle	92						

Db	654	GCATTTGTGCTTGCATTTTCCCTTGGCATGGGAGCCATCCCATGGATCATAATGTCTGAG	713
Qy	464	IleLeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeu	483
Db	714	ATCCTTCTCTGTTAACATCAAGAGTCTGCCGGGAAGCACCCCCACCTCGCAAACTGGATG	773
Qy	484	ThrSerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThr	503
Db	774	ACGTCTGGCTCATCACGATACCGGAGCTTGATGCTAAACTGGAGCAACGGAGGAACC	833
Qy	504	PheLeuSerTyrMetValValSerAlaPheThrIleValPheValValLeuTrpValPro	523
Db	834	TTGCTATA-TTCGGCGGGTGTCCATGGGCACCCCTCTCTTCGTGCTGTGCGTGCCG	892
Qy	524	GluThrLysGly***AsnSerArgGlyAspThrIlePheValSerLeuSer	540
Db	893	GAGACCAAGGGAAG-AACATTCGAGGAAA-TCGGTTTTTCGTTCCGCTGAAC	941
RESULT 8			
AV939950			
LOCUS	AV939950	674 bp	mRNA linear EST 18-JAN-2002
DEFINITION	AV939950 K. Sato unpublished cDNA library, strain H602 adult, heading stage top three leaves Hordeum vulgare subsp. spontaneum cDNA clone bah24ml2 5', mRNA sequence.		
ACCESSION	AV939950		
VERSION	AV939950.1	GI:18235747	
KEYWORDS	EST.		
SOURCE	Hordeum vulgare subsp. spontaneum		
ORGANISM	Hordeum vulgare subsp. spontaneum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Hordeum.		
AUTHORS	1 (bases 1 to 674)		
TITLE	Sato, K., Saisho, D. and Takeda, K.		
JOURNAL	Barley EST sequencing project in NIG and Okayama Univ Unpublished (2002)		
COMMENT	Contact: Tadasu Shin-i Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.		
FEATURES	Location/Qualifiers		
source	1..674		
	/organism="Hordeum vulgare subsp. spontaneum"		
	/mol_type="mRNA"		
	/strain="H602"		
	/sub_species="spontaneum"		
	/db_xref="taxon:77009"		
	/clone="bah24ml2"		
	/tissue_type="top three leaves"		
	/dev_stage="adult, heading stage"		
	/clone_lib="K. Sato unpublished cDNA library, strain H602 adult, heading stage top three leaves"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.44e-91	Length:	674
Score:	984.00	Matches:	193
Percent Similarity:	91.93%	Conservative:	12
Best Local Similarity:	86.55%	Mismatches:	17
Query Match:	35.41%	Indels:	1
DB:	1	Gaps:	0
US-10-051-909-36 (1-553) x AV939950 (1-674)			
Qy	53	GlusSerGlySerAspHisAspGlyValLeuArgArgPro-LeuLeuAsnThrGlySerTr	72
Db	3	GAGAGCGGAGCGGACCATGACACCGGNAAGCAGCCGCTGCTGGTCAACACGGGGAGCTG	62
Qy	72	pTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaValle	92



Db	63	GTACCGGATGGGTGCGGCCAGTCCAGCCTCACGGCGGCACCTCTCTCCATGGCCATCAT	122
QY	92	uArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIleGl	112
Db	123	GCGGAGTCCCACGTCTCCGCCTTCTCTGACCACATGATCGTCGCGTCCGCCCATCCA	182
QY	112	nPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeuAs	132
Db	183	GTTCGGCTTCACCGCGGCTTCTCCTCCCCACCCAGGACGCCATCATCGCGACCTCAA	242
QY	132	nLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetValGl	152
Db	243	CCTCTCCATCTCCAGATTCTCCGTGTTCCGCTCGTCTCCAACTGCGCGGCCATGGTCGG	302
QY	152	yAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIleAl	172
Db	303	GGCCATCGCCAGCGCCAGATGGCGCGAGCACATTGGCGCGCAAAGGGTCACTGATTTGC	362
QY	172	aAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLe	192
Db	363	TGCTATTCTTAACATCATCGGCTGGCTGGCCATCTCTTCGCAAAAGACACTTCTTTTCT	422
QY	192	uTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValProVa	212
Db	423	GTATATGGGACGATTGCTCGAAGGATTGGTGTGGTGTGCATATCTACACGTGCCAGT	482
QY	212	lTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLe	232
Db	483	ATACATAGCAGAGATTCTCTCTCAGAACATGAGAGCGGCTCTAGGCTCTGTGAACAGTT	542
QY	232	uSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArgLe	252
Db	543	GTCGGTAACGATTGGTATCGTGTGGCTTACATTCTCGGCATGTTTGTTCCTTGGAGGAT	602
QY	252	uLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIlePr	272
Db	603	GCTTGCAGTGATAGGAATCTTGCCCATGCACAATATTGATACCTGGTCTGTTCTTCATTCC	662
QY	272	oGluSer	274
Db	663	CGAATCT	669
RESULT 9			
CA262285			
LOCUS			
DEFINITION	CA262285	714 bp mRNA linear EST 26-SEP-2003	
		SCEQLB2019E02.9 LB2 Saccharum officinarum cDNA clone SCEQLB2019E02	
		5', mRNA sequence.	
ACCESSION	CA262285		
VERSION	CA262285.1	GI:35954083	
KEYWORDS	EST.		
SOURCE	Saccharum officinarum		
ORGANISM	Saccharum officinarum		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD		
	clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum		
	complex.		
REFERENCE	1	(bases 1 to 714)	
AUTHORS	Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.		
TITLE	The libraries that made SUCEST		
JOURNAL	Genet. Mol. Biol. 24 (1-4), 1-7 (2001)		
COMMENT	Contact: Arruda P		
	Centro de Biologia Molecular e Engenharia Genetica		
	Universidade Estadual de Campinas		
	Caixa Postal 6010, 13083-970, Campinas SP, Brazil		
	Tel: 55 19 3788 1137		
	Fax: 55 19 3788 1089		
	Email: parruda@unicamp.br		
	Clone distribution: clone distribution information can be found		
	through the Brazilian Clone Collection Center (BCCC) at		
	http://www.bccccenter.fcav.unesp.br		
	Plate: 019 row: E column: 02		
	Seq primer: T7 Promoter Primer.		
	Location/Qualifiers		
FEATURES			

		source	
		1..714	
		/organism="Saccharum officinarum"	
		/mol_type="mRNA"	
		/db_xref="taxon:4547"	
		/clone="SCEQLB2019E02"	
		/lab_host="DH10B"	
		/clone_lib="LB2"	
		/note="Organ: Lateral buds from plants adult plants	
		growing in greenhouse; Vector: pSport1; Site_1: SalI;	
		Site_2: NotI; An unidirectional cDNA library generated	
		from [lateral buds from plants adult plants growing in	
		greenhouse]. cDNA was prepared from polyA+ mRNA using	
		SuperScript plasmid System Kit (Invitrogen). The	
		double-strand cDNAs were fractionated in a sepharose	
		CL-2B 40cm-columns and fragments sizing between 0.8 and	
		1.5 Kb were directionally cloned into the vector. Details	
		of each source of RNA and library construction can be	
		obtained at http://sucest.lad.ic.unicamp.br/public"	
		ORIGIN	
		Alignment Scores:	
		Pred. No.: 1.81e-90 Length: 714	
		Score: 976.00 Matches: 188	
		Percent Similarity: 90.52% Conservative: 22	
		Best Local Similarity: 81.03% Mismatches: 22	
		Query Match: 35.12% Indels: 0	
		DB: 6 Gaps: 0	
		US-10-051-909-36 (1-553) x CA262285 (1-714)	
QY	92	LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle	111
Db	15	CTCCGCGACGTGCGCCATCTCGGCCACGTTCTGACAGCTCATTTGTGGCTCGAGGTCCC	74
QY	112	GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu	131
Db	75	CAGTTCGGTTTCACCAAGCGGTACTCTCTCCCGACGAGGAGGAGGAGGAGGAGGAG	134
QY	132	AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal	151
Db	135	GGCCTCTCCCTCTCCGAGTTCCTCGCTCTTCGGATCGTTATCAACGCTAGGGGCAAT	194
QY	152	GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle	171
Db	195	GGTGCCATCTCCAGTGGGCAACTTGCAGAGTATATCGGTGCGTGTGTAATATCTTAT	254
QY	172	AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe	191
Db	255	GCTGCAATTCCAAACATATAATTGGTGGCTCGCGATATCATTCGCAAAAGATTCTCT	314
QY	192	LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro	211
Db	315	TTGTTTATGGTTCGGCTGCTAGAAGGATTTGGAGTCGGTGTGTAATATCTTATACAG	374
QY	212	ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln	231
Db	375	GTTTATATGCAGAAATCGCTCCTCAAGATCAGAGGGGAGTTCTTGGTTCTGTCAAT	434
QY	232	LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTrpArg	251
Db	435	CTCTCCGTCACAATTGGTATATTGCTTGCTTGCCTACCTGTTTGGCATGTTTGTTC	494
QY	252	LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle	271
Db	495	ATTCTTGCTGTCTAGGCGTTTACCTTGTTTCAATACTGATTCCTGGACTGTTCTTT	554
QY	272	ProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCysGluThrSerLeu	291
Db	555	CCTGAATCCCCAAGGTGGCTGGCAAAAATGGGGAAGATGGAGGATTTTGAATATT	614
QY	292	GlnValLeuArgGlyPheGluThrAspIleThrThrGluValAsnAspIleLysArgAla	311
Db	615	CAAAGTTCTGCGAGGATTTCAGACGGACATCACACAGAAAGTAAATGAAATAAGAG	674

QY 312 ValAlaSerSerSerLysArgThrThrIleSerPhe 323  
|||||  
Db 675 GTAGCATCATCAAGGAAGAGGACAACCATATGTTT 710

RESULT 10  
AV913420  
LOCUS AV913420 681 bp mRNA linear EST 18-JAN-2002  
DEFINITION AV913420 K. Sato unpublished cDNA library, cv. Haruna Nijo  
germination shoots Hordeum vulgare subsp. vulgare cDNA clone  
bags22a03 5', mRNA sequence.

ACCESSION AV913420  
VERSION AV913420.1 GI:18209197  
KEYWORDS EST.  
SOURCE Hordeum vulgare subsp. vulgare  
ORGANISM Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Pooideae; Triticeae; Hordeum.  
1 (bases 1 to 681)  
Sato,K., Saisho,D. and Takeda,K.  
Barley EST sequencing project in NIG and Okayama Univ  
Unpublished (2002)  
Contact: Tadasu Shin-i  
Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

FEATURES  
source Location/Qualifiers  
1..681  
/organism="Hordeum vulgare subsp. vulgare"  
/mol\_type="mRNA"  
/cultivar="Haruna Nijo"  
/sub\_species="vulgare"  
/db\_xref="taxon:112509"  
/clone="bags22a03"  
/tissue\_type="shoots"  
/dev\_stage="germination"  
/clone\_lib="K. Sato unpublished cDNA library, cv. Haruna  
Nijo germination shoots"

ORIGIN  
Alignment Scores:  
Pred. No.: 2.14e-90 Length: 681  
Score: 975.00 Matches: 191  
Percent Similarity: 91.89% Conservative: 13  
Best Local Similarity: 86.04% Mismatches: 16  
Query Match: 35.08% Indels: 2  
DB: 1 Gaps: 2

US-10-051-909-36 (1-553) x AV913420 (1-681)

QY 53 GluSerGlySerAspHisAspGlyValLeuArgArgProLeuLeu---AsnThrGlySer 71  
|||  
Db 18 GAGGAGAGAGAGACCATGAC---ACGCGGAAGCAGCCGCTGCTGGTCAACACGGGGAGC 74

QY 72 TrpTyrArgMetSerSerArgGlnSerSerPheAlaProGlyThrSerSerMetAlaVal 91  
|||||  
Db 75 TGGTACCGGATGGGGTCGGCCAGTCCAGCCTCAGGCGGGCACCTCTCTCCATGGCCATC 134

QY 92 LeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAlaLeuGlyProIle 111  
:::  
Db 135 ATGCGGGAGTCCACGCTCTCCGCCTTCTCTGCACCATGATCGTCGCGTCGGCCCCCATC 194

QY 112 GlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMetValArgAspLeu 131  
|||||  
Db 195 CAGTTCGGCTTCACCGCGGCTTCTCTCCCCACCCAGGACGCCATCATCCGCGACCTC 254

QY 132 AsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnValGlyGlyMetVal 151  
|||||  
Db 255 AACCTCTCCATCTCCGAGTTCTCCGTGTTCCGCTCGCTGTCCAACGTCGGCGCCATGGTC 314

QY 152 GlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGlySerLeuMetIle 171  
|||||  
Db 315 GGCGCCATCGCCAGCGGCAGATGGCCGAGCACATTGGCCGCAAAGGTCACTGATGATT 374

QY 172 AlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLysAspAlaSerPhe 191  
|||||  
Db 375 GCTGCTATTCTTAACATCATCGGCTGGCTGGCCATCTCCTTCGAAAAGACACCTCTTTT 434

QY 192 LeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSerTyrThrValPro 211  
|||||  
Db 435 CTGTATATGGGACGATTGCTCGAAGGATTGGTGTGGTGTCATATCTACACGGTGCCA 494

QY 212 ValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGlySerValAsnGln 231  
|||||  
Db 495 GTATACATAGCAGAGATTCTCTCTCAGAACATGAGAGGCGCTCTAGGCTCTGTGAACCA 554

QY 232 LeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPheIleProTyrArg 251  
|||||  
Db 555 TTGTCGGTAACGATTGGTATCGTGTGGCCTACATTCTCGGCATGTTTGTTCCTTGAGG 614

QY 252 LeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGlyLeuPhePheIle 271  
:::  
Db 615 ATGCTTGCAGTGATAGGAATCTTGCCATGCACAATATTGATACCTGGTCTGTTCTTCATT 674

QY 272 ProGlu 273  
|||||  
Db 675 CCCGAA 680

RESULT 11  
CR291514  
LOCUS CR291514 718 bp mRNA linear EST 27-FEB-2004  
DEFINITION CR291514 Oryza sativa library (Han B) Oryza sativa cDNA clone  
y745e11p5, mRNA sequence.

ACCESSION CR291514  
VERSION CR291514.1 GI:44678080  
KEYWORDS EST.  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.  
1 (bases 1 to 718)  
Han,B., Feng,Q., Huang,Y.C., Ying,K., Li,Y., Guan,J.P., Zhu,J.J.,  
Zhao,Q., Hu,X., Liu,Y.L., Mu,J., Yu,Z., Chen,L., Fan,D.L.,  
Weng,Q.J., Zhang,L., Lu,Y.Q., Yu,S.L., Liu,X.H., Lu,T.T.,  
Zhang,Y.J., Lu,Y., Li,C., Li,T., Zhang,Y., Hu,H., Jia,P.X.,  
Zhang,L., Lan,L.F., Chen,W., Wu,S.A. and Xue,Y.B.  
Rice cDNA EST clone  
Unpublished (2003)  
Contact: Han Bin  
National Center for Gene Research  
Chinese Academy of Sciences  
500# Cao Bao Road, Shanghai 200233, China  
Email: bhan@ncgr.ac.cn  
Clone requests: bhan@ncgr.ac.cn  
This is rice cdna est clone  
Web site: http://www.ncgr.ac.cn.

FEATURES  
source Location/Qualifiers  
1..718  
/organism="Oryza sativa"  
/mol\_type="mRNA"  
/db\_xref="taxon:4530"  
/clone="y745e11p5"  
/clone\_lib="Oryza sativa library (Han B)"

ORIGIN  
Alignment Scores:  
Pred. No.: 4.62e-89 Length: 718  
Score: 962.50 Matches: 198  
Percent Similarity: 93.42% Conservative: 15  
Best Local Similarity: 86.84% Mismatches: 15  
Query Match: 34.63% Indels: 3

Fri Apr 15 06:53:46 2005

DB:		7	Gaps:	0
US-10-051-909-36 (1-553) x CR291514 (1-718)				
QY	160	AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly	179	
Db	8	CGGAGCTACATTGGGCGCAAGGGTCA-TTGATGATTGCTGCAATTCCAAACATCATTTGGT	66	
QY	180	TTrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu	199	
Db	67	TGGCTTGCCATCTCCTTTGCAAAAGACTCATCGTTTCTTTATATGGGACGATTGCTCGAG	126	
QY	200	GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro	219	
Db	127	GGGTTTGGTGTGGTGTCTATACGGTGCCAGTTTACATAGCAGAAATATCACCT	186	
QY	220	GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe	239	
Db	187	CAAAACATGAGAGGTGCTCTTGGCTCAGTGAATCAGTTATCTGTAAACCGTTGGTATATTG	246	
QY	240	LeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaValIleGlyAlaLeu	259	
Db	247	TTGGCATATTTTGCTCGGCATGTTTGTTCCTTGGAGGCTTCTGTGTAATAGGAATCTTG	306	
QY	260	ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTTrpLeuAla	279	
Db	307	CCTTGCACTGTGTGATACCTGGCCTATTTCTTCAAGTTCTGAGGGATTTTGAGACT	366	
QY	280	LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr	299	
Db	367	AAGATGAACATGATGATGATTTTGAGACTTCTTTACAAGTTCTGAGGGGATTTTGAGACT	426	
QY	300	AspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThr	319	
Db	427	GACATCAGCGCGGAGTGAATGATATAAAGAGAGCAGTAGCTTCAGCAAAACAAAGGACA	486	
QY	320	ThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIle	339	
Db	487	ACGATCCGTTTTCAGAAATTAACACAGAGAAATACCGCACACCCCTAATACTAGGAATT	546	
QY	340	GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer	359	
Db	547	GGCCTACTTGTACTGCAACAGGTAAGTGAATCACCGGAATATTGTTTATGCAGGTAGC	606	
QY	360	IlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle	379	
Db	607	ATCTTCCAAGCAGCAGGTCTCACAAACAGTGACTTGGGTACATGTGCACCTGGGGGT-ATC	665	
QY	380	GlnValLeuAlaThrGlyValThr	387	
Db	666	CAGGTTCT-GCTACAGGAGTTACA	688	
RESULT 12				
AJ794429				
LOCUS				
DEFINITION				
018_3_05_d17, mRNA sequence.				
AJ794429				
VERSION				
AJ794429.1				
GI:51109757				
KEYWORDS				
SOURCE				
ORGANISM				
Antirrhinum majus (snapdragon)				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
asterids; lamids; Lamiales; Plantaginaceae; Antirrhineae;				
Antirrhinum.				
1 (bases 1 to 762)				
Zachgo,S., Stueber,K., Saedler,H., Sommer,H. and Schwarz-Sommer,Z.				
AUTHORS				
TITLE				
Antirrhinum EST collection				
JOURNAL				
Unpublished (2003)				
COMMENT				
Contact: Schwarz-Sommer Z				
Molekulare Pflanzen-genetik				
MPI fuer Zuechtungs-forschung				
FEATURES				
source				
1..762				
Carl-von-Linne Weg 10, D-50829, Germany.				
Location/Qualifiers				
/organism="Antirrhinum majus"				
/mol_type="mRNA"				
/db_xref="taxon:4151"				
/clone="018_3_05_d17"				
/tissue_type="whole plant"				
/clone_lib="Antirrhinum majus whole plant"				
ORIGIN				
Alignment Scores:				
Pred. No.: 5.08e-89 Length: 762				
Score: 962.50 Matches: 189				
Percent Similarity: 86.00% Conservative: 26				
Best Local Similarity: 75.60% Mismatches: 34				
Query Match: 34.63% Indels: 1				
DB: 1 Gaps: 1				
US-10-051-909-36 (1-553) x AJ794429 (1-762)				
QY	88	SerMetAlaValLeuArgGluSerHisValSerAlaPheLeuCysThrLeuIleValAla	107	
Db	13	TCTTCTCAAGCTTTTAGGGACAGTTCTGTTCCTGTGTGGCTTGTGTAAATTGTGGCT	72	
QY	108	LeuGlyProIleGlnPheGlyPheThrSerGlyPheSerSerProThrGlnAspAlaMet	127	
Db	73	TTGGTCTCTATCCAGTTTGGTTCACAAGTGGTATCTTCCACGACTCAAACGTCTATT	132	
QY	128	ValArgAspLeuAsnLeuSerIleSerGluPheSerAlaPheGlySerLeuSerAsnVal	147	
Db	133	GTCAATGATCTTAAACTCACAGTTTCCACAGTTCTCGCTATTGTGGTCTTGTGGAACGTT	192	
QY	148	GlyGlyMetValGlyAlaIleAlaSerGlyGlnMetAlaGluTyrIleGlyArgLysGly	167	
Db	193	GGAGCCATGGTTGGAGCAATAGCTAGTGGTCAGATTGCTGAGTACATAGGAGAGAAAGGG	252	
QY	168	SerLeuMetIleAlaAlaIleProAsnIleIleGlyTrpLeuAlaIleSerPheAlaLys	187	
Db	253	TCTTTAATGATCGCTGCCATACCTAAATATCATTTGGTTGGCTTGGCGTCCGCAATACTCT	312	
QY	188	AspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGlyPheGlyValGlyIleIleSer	207	
Db	313	GACATCTCATTTCTGTACATGGGAAGATTGTTGGAAGGATTTGGCGTCGGCATAAATCTCT	372	
QY	208	TyrThrValProValTyrIleAlaGluIleSerProGlnAsnMetArgGlyAlaLeuGly	227	
Db	373	TATACTGTGCCTGTGTATATAGCTGAGATAGCACCTGAAAATCTAAGGGGGGCTCTGGGG	432	
QY	228	SerValAsnGlnLeuSerValThrPheGlyIlePheLeuAlaTyrLeuLeuGlyMetPhe	247	
Db	433	TCAGTAAACCCAGCTCTCTGTGACGATCGGGATCATGTAGCATATTTACTTTGGACTATT	492	
QY	248	IleProTrpArgLeuLeuAlaValIleGlyAlaLeuProCysThrMetLeuIleProGly	267	
Db	493	TGCAATTGGAGATTGCTTGTGTTCTAGGAATATTGCCCTTGTCTGATATTGATACCTGGC	552	
QY	268	LeuPhePheIleProGluSerProArgTrpLeuAlaLysMetAsnLeuThrGluAspCys	287	
Db	553	CTCTTTTTCATCCAGAAATCTCCTAGATGGTTGGCCAAATGGGGATGACAGAGAATTT	612	
QY	288	GluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrGluValAsnAsp	307	
Db	613	GAAGCCTCTCTTCAAGTTCTTCGGGGGTTTGACGCTGATATTGCAGTTGAAGTATCTGAG	672	
QY	308	IleLysArgAlaValAlaSerSerSerLysArgThr--ThrIleSerPheGlnGluLeu	326	
Db	673	ATAAAGAAATCTGTGCTTCAACGAGCAGACGACGGGAGCAATCCGCTTTGCCGATCTT	732	
QY	327	AsnGlnLysLysTyrArgThrProLeuLeu	336	
Db	733	AAGTTGAAAAGATACTGTTTACCGTTGATG	762	



RESULT 13	
CD8711121	CD8711121
LOCUS	AZO2.117H06F010207 AZO2 Triticum aestivum cDNA clone AZO2117H06,
DEFINITION	mRNA sequence.
ACCESSION	CD8711121
VERSION	CD8711121.1 GI:32554937
KEYWORDS	EST.
SOURCE	Triticum aestivum (bread wheat)
ORGANISM	Triticum aestivum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
	Poideae; Triticeae; Triticum.
REFERENCE	1 (bases 1 to 701)
AUTHORS	Genoplante.
JOURNAL	Genoplante, a major partnership french program in plant genomics
COMMENT	Unpublished (2003)
	Contact: Genoplante
	Genoplante
	93, rue Henri Rochefort 91025 EVRY CEDEX France
	Tel: 33 1 69 47 54 00
	Fax: 33 1 69 47 54 10
	This sequence has been generated in the framework of the french
	plant genomics programme 'Genoplante' ( <a href="http://www.genoplante.com">http://www.genoplante.com</a>
	and <a href="http://genoplante-info.infobiogen.fr">http://genoplante-info.infobiogen.fr</a> ).
FEATURES	Location/Qualifiers
source	1..701
	/organism="Triticum aestivum"
	/mol_type="mRNA"
	/cultivar="recital"
	/db_xref="taxon:4565"
	/clone="AZO2117H06"
	/tissue type="root"
	/clone_lib="AZO2"
ORIGIN	
Alignment Scores:	
Pred. No.:	6.55e-86
Score:	932.00
Percent Similarity:	90.48%
Best Local Similarity:	78.79%
Query Match:	33.54%
DB:	6
	Gaps: 0
US-10-051-909-36 (1-553) x CD8711121 (1-701)	
QY	285 GluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAspIleThrThrGlu 304
Db	
	7 GAGGACTTTGAAACCTCTCTACAAGTTCTGAGGGGATTTGAGACTGACATCACCTCAGAA 66
QY	305 ValAsnAspIleLysArgAlaValAlaSerSerLysArgThrThrIleSerPheGln 324
Db	
	67 GTGAATGATATAAAGAGAGCAGTAATATCAGCAAAACAAAGGGCGGCGATCCGTTTCCAA 126
QY	325 GluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIleGlyLeuLeuValLeu 344
Db	
	127 GAGTTAAACCAAAAGAAATTCCGCATGCCCTGATTCTAGGAATTGGCCTGCTTGTCTA 186
QY	345 GlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIlePheLysAlaAla 364
Db	
	187 CAACAGCTAAGCGGAATCAACGCTATACTGTTCTATGCAAGTAGCATCTTCAAAGCTGCA 246
QY	365 GlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGlnValLeuAlaThr 384
Db	:    :    :    :    :    :    :    :    :    :    :    :    :    :
	247 GGTATTACAACAGTGACTTGGCCACATCTGGACTTGGAGGTATTGAGGTCTTGCCACT 306
QY	385 GlyValThrThrTrpLeuLeuAspArgAlaGlyArgGileLeuLeuIleIleSerThr 404
Db	
	307 CTAGTTACAACCTGGTTACTAGACAGGGCTGGCCGGGTATCCTACTCATATATCTTCT 366
QY	405 SerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeuLysAspAsnIle 424
Db	:    :    :    :    :    :    :    :    :    :    :    :    :    :
	367 GCTGGGATGACTATAAGCCTTCTTGCGGTGCCGTCATATTTTATCAAGGACACTGTT 426
QY	425 SerGlnAspSerAsnSerTyrTyrIleLeuThrMetIleSerLeuValGlyIleValSer 444
Db	
	427 TCACAAGACTCTCACATGTATTACATATTGAGCATGGTCTCCTTGCTGTGTTGGCT 486
QY	445 PheValIleThrPheSerPheGlyMetGlyAlaIleProTrpLeuMetMetSerGluIle 464
Db	:    :    :    :    :    :    :    :    :    :    :    :    :    :
	487 TATGTTATCGCCTTCTCCTTCGGTATGGGGCCATTCCATGGGTCAATAATGTCTGAGATT 546
QY	465 LeuProValSerIleLysSerLeuGlyGlySerIleAlaThrLeuAlaAsnTrpLeuThr 484
Db	
	547 CTCCCGTGAGCATCAAGAGTCTCGCGGGAAGCTTCGCGACGCTCGCCAACCTGGCTGACT 606
QY	485 SerPheAlaIleThrMetThrThrAsnLeuMetLeuThrTrpSerValGlyGlyThrPhe 504
Db	
	607 TCCTTTGGGATAACGATGACAGCAAACTTGCTGCTCAGCTGGAGTGTGGAGGTACATTT 666
QY	505 LeuSerTyrMetValValSerAlaPheThrIle 515
Db	:    :    :    :    :    :    :    :    :    :    :    :    :    :
	667 GTGTTTACATGCTCGTGAGCGGCTTCACGCTC 699
RESULT 14	
CNSOA8HU	
LOCUS	1675 bp mRNA linear HTC 06-FEB-2004
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone
	GSLTLS242D01 of Adult vegetative tissue of strain col-0 of
	Arabidopsis thaliana (thale cress).
ACCESSION	BX819759
VERSION	BX819759.1 GI:42467369
KEYWORDS	HTC; GSLT cDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE	1 (bases 1 to 1675)
AUTHORS	Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
	Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
	Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
	A Combined Approach to Evaluate and Improve Arabidopsis Genome
	Annotation
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1675)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
	BP 191 91006 EVRY cedex - FRANCE (E-mail : <a href="mailto:segref@genoscope.cns.fr">segref@genoscope.cns.fr</a>
	- Web : <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a> )
COMMENT	The sequences are based on single pass reads.
	Life Technologies (a division of Invitrogen) members carried out
	full-length libraries construction : Temple G.
	Genoscope members carried out sequencing and annotation : Castelli
	V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
	Schachter V., Weissenbach J., Salanoubat M.
	URGV INRA : Clepet C., Caboche M.
	Annotation is based on the June 2003 version of the Arabidopsis
	genome released by MIPS (Munich Information center for Protein
	Sequences). 5 prime and 3 prime are assembled with Phrap.
	<a href="http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full">http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full</a>
	length
	<a href="http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis">http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis</a> .
FEATURES	Location/Qualifiers
source	1..1675
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/strain="Col-0"
	/db_xref="taxon:3702"
	/clone="GSLTLS242D01"
	/tissue_type="Adult vegetative tissue"
	/plasmid="pCMVSPORT_6"
	1..1675
gene	/gene="At2g48020"

ORIGIN

Alignment Scores:	3.74e-85	Length:	1675
Pred. No.:	930.50	Matches:	180
Score:	65.73%	Conservative:	102
Percent Similarity:	41.96%	Mismatches:	141
Best Local Similarity:	33.48%	Indels:	6
Query Match:	3	Gaps:	2

US-10-051-909-36 (1-553) x CNS0A8HU (1-1675)

QY	100	PheLeuCysThrLeuIleValAlaLeuGlyProIleGlnPheGlyPheThrSerGlyPhe	119
Db	201	TATCTCAGCACATTCGTTGGCGTCTGTGGTTCTTTTGGCTCGGTTCTTGTGGGGGATAC	260
QY	120	SerSerProThrGlnAspAlaMetValArgAspLeuAsnLeuSerIleSerGluPheSer	139
Db	261	TCCTCACCTGCTCAGGCTGCAATTAGGAATGACCTTTTCATGACTATAGCTGAGTTTCA	320
QY	140	AlaPheGlySerLeuSerAsnValGlyGlyMetValGlyAlaIleAlaSerGlyGlnMet	159
Db	321	CTCTTTGGTTCTTTACTAACTTTTGGCGCAATGATCGGTGCTATAAACAAGCGGGCTATA	380
QY	160	AlaGluTyrIleGlyArgLysGlySerLeuMetIleAlaAlaIleProAsnIleIleGly	179
Db	381	GCTGATTGATGGAAGAAAGGGCGATGAGAGTTTCTCTGTCATTTTGTGAGTCGGG	440
QY	180	TrpLeuAlaIleSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGlu	199
Db	441	TGGCTAGCAATCATCTTTTGCCAAAGGGAGTAGTGGCTCTGGACCTTGGAAAGACTTGGCAACG	500
QY	200	GlyPheGlyValGlyIleIleSerTyrThrValProValTyrIleAlaGluIleSerPro	219
Db	501	GGATATGGAATGGGAGCATTTTCCATGTGGTGCCAAATCTTTATAGCAGAAATTGCACCT	560
QY	220	GlnAsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePhe	239
Db	561	AAAACTTTCAGAGGGGCTCTAAACCACACTGACCTGAACAGATTTCTGATCTGCACCTGGAGTGTG	620
QY	240	LeuAlaTyrLeuLeuGlyMetPheIleProTyrArgLeuLeuAlaIleGlyAlaLeu	259
Db	621	GTTTCCTTCATCATAGGCACACTAGTAGCAGCTGGAGAGTCTTGGCATTATATAGGAATCATC	680
QY	260	ProCysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAla	279
Db	681	CCATGCGCTGCCCTCCTCTCTTGGCCCTCTTTTATCCCTGAGTCTCCAAGATGGCTGGCA	740
QY	280	LysMetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThr	299
Db	741	AAAGTGGGCGTGATACGGAGTTTGAAGCTGCACCTGAGGAAGCTCCGTGGGAAGAGGCT	800
QY	300	AspIleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThr	319
Db	801	GATAITTCGGAGGAGGCAGCAGAGATCCAGGATTATATCGAAACTCTGGAAAGGCTACCG	860
QY	320	ThrIleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuLeuGlyIle	339
Db	861	AAAGCCAAAGATGCTGGATTGTTTTCAGAGGAGATACATACGCTCTGTCTTATAGCTTTTC	920
QY	340	GlyLeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSer	359
Db	921	GGGTTGATGGTGTTCAGCAGCTTTGGAGGAATCAACGGAATATGTTTCTACACAAGCTCG	980
QY	360	IlePheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIle	379
Db	981	ATATTTTGACCAAGCAGGTTTTTCCC---ACAAGACTTGGGATGATAATATATGCTGTCTT	1037
QY	380	GlnValLeu-AlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLe	399
Db	1038	CAGGTGGTAAATCACTCGCGCTTAATGCAACCGATAGTTGACAGAGCCGGAAGAAACCAT	1097
QY	399	uLeuIleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValValPhePh	419

Pred. No.: 4.05e-85 Length: 769  
Score: 925.00 Matches: 177  
Percent Similarity: 85.66% Conservative: 38  
Best Local Similarity: 70.52% Mismatches: 36  
Query Match: 33.29% Indels: 0  
DB: 4 Gaps: 0

US-10-051-909-36 (1-553) x BI933646 (1-769)

Qy	181	LeuAlaileSerPheAlaLysAspAlaSerPheLeuTyrMetGlyArgLeuLeuGluGly	200
Db	1	CTTTCCATCTCGTTGCCAAGATCTGTCCTTCTATACATGGGAAGATTATTGGAAGGT	60
Qy	201	PheGlyValGlyIleileSerTyrThrValProValTyrIleAlaGluileSerProGln	220
Db	61	TTTGGAGTCGCGATAATATCTTACACGGTTCCTGTATATATTGCCGAGATAGCACCTCAG	120
Qy	221	AsnMetArgGlyAlaLeuGlySerValAsnGlnLeuSerValThrPheGlyIlePheLeu	240
Db	121	AACCTGAGAGGGGCCCTGGGGTCAGTTAACAGCTCTCTGTTACAAATTGGGATCATGTTG	180
Qy	241	AlaTyrLeuLeuGlyMetPheIleProTrpArgLeuLeuAlaValIleGlyAlaLeuPro	260
Db	181	GCTTATTGCTAGGACTTTTGTAAATTGGAGAGTGCTCGCTTTTCTTGGAAACATTGCC	240
Qy	261	CysThrMetLeuIleProGlyLeuPhePheIleProGluSerProArgTrpLeuAlaLys	280
Db	241	TGCCTTGCATTGATACCTGGCCTATTTTTCATCCAGAAATCTCCTCGGTGGTTGGCCAAG	300
Qy	281	MetAsnLeuThrGluAspCysGluThrSerLeuGlnValLeuArgGlyPheGluThrAsp	300
Db	301	ATGGGTCTGCAGATGATTTTGAACCTCTTTGCAAGTTCTCCGAGGGTTCGATGCTGAC	360
Qy	301	IleThrThrGluValAsnAspIleLysArgAlaValAlaSerSerSerLysArgThrThr	320
Db	361	ATTTCCGTCGAAGTAAATGAAATTAAGAGGGCTGTAGCATCCACAGCCGAAAGTCAACA	420
Qy	321	IleSerPheGlnGluLeuAsnGlnLysLysTyrArgThrProLeuLeuGlyIleGly	340
Db	421	ATACGTTTTTGCAGATCTCAACAAAGAGATATTGGCTGCCTCTCATGATAGGCATTGGA	480
Qy	341	LeuLeuValLeuGlnAsnLeuSerGlyIleAsnGlyValLeuPheTyrAlaSerSerIle	360
Db	481	CTGCTTGTCCTACAACTCACTCAGCGGAACCAATGGTGTGATCTTCTATTCCAGTAACATT	540
Qy	361	PheLysAlaAlaGlyValThrAsnSerAspLeuAlaThrCysSerLeuGlyAlaIleGln	380
Db	541	TTCCCTATCGGCCGGGATTTCTTCAAGTGATGCTGCAACTTTAGGTTTGGTGCTATCCAG	600
Qy	381	ValLeuAlaThrGlyValThrThrTrpLeuLeuAspArgAlaGlyArgArgIleLeuLeu	400
Db	601	GTGGTTGCCACTGCTGTTTCTACATGGCTGGTGATATAAACTGGCCGTAGGCTTTTACTG	660
Qy	401	IleIleSerThrSerGlyMetThrLeuCysLeuLeuAlaValSerValPhePheLeu	420
Db	661	ATTGTCTCGTCAGCTGGAATGGCTGGTAGTCTCCTTATTGTTTCCATTGCATTCTTTGGT	720
Qy	421	LysAspAsnIleSerGlnAspSerAsnSerTyr	431
Db	721	AAGGATTTCTAGATGAGGATTCTACCTTCTAT	753



*This Page Blank (uspro)*